Spatially explicit capture-recapture for bear researchers and managers

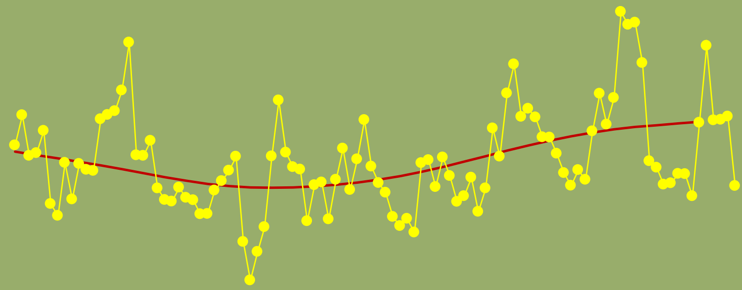
Murray Efford
Western Black Bear Workshop
Coeur d’Alene ID May 2012

Morning session

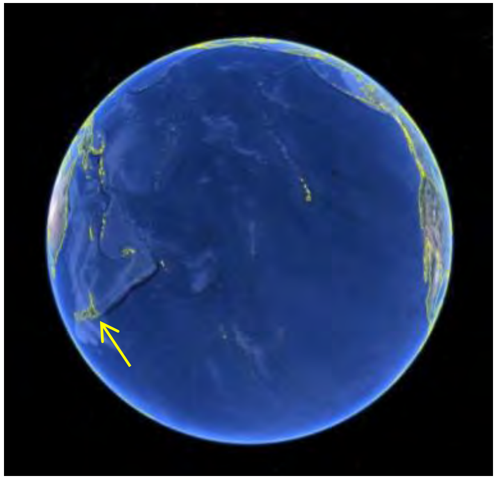
1. Introduction

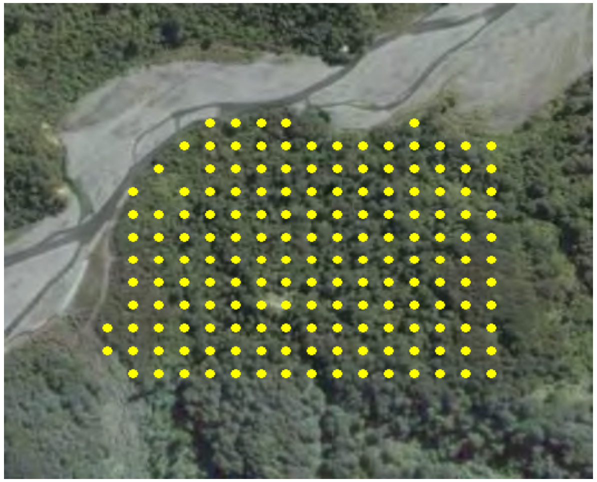
2. DENSITY software
    - Interface
    - Data
    - Conventional analyses
    - Simple spatial analysis: GSM black bears

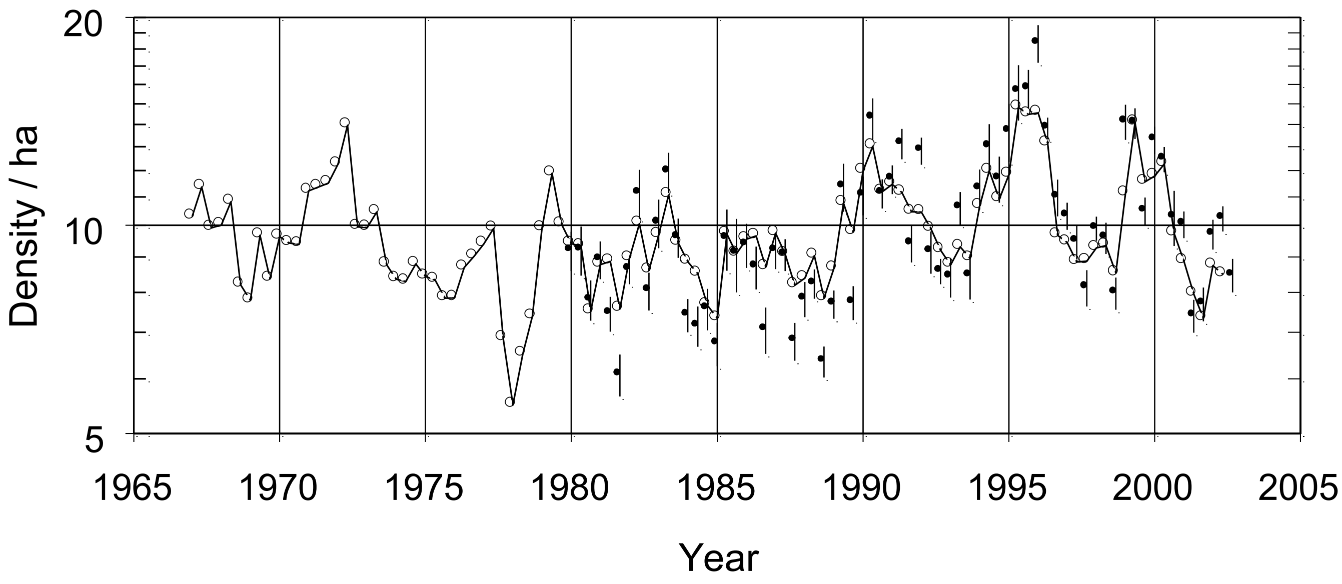
3. Key concepts
    - Detector types
    - Buffers, habitat masks, and the 'region of integration'
    - Maximum likelihood



Long-term change in population density of brushtail possums (*Trichosurus vulpecula*)

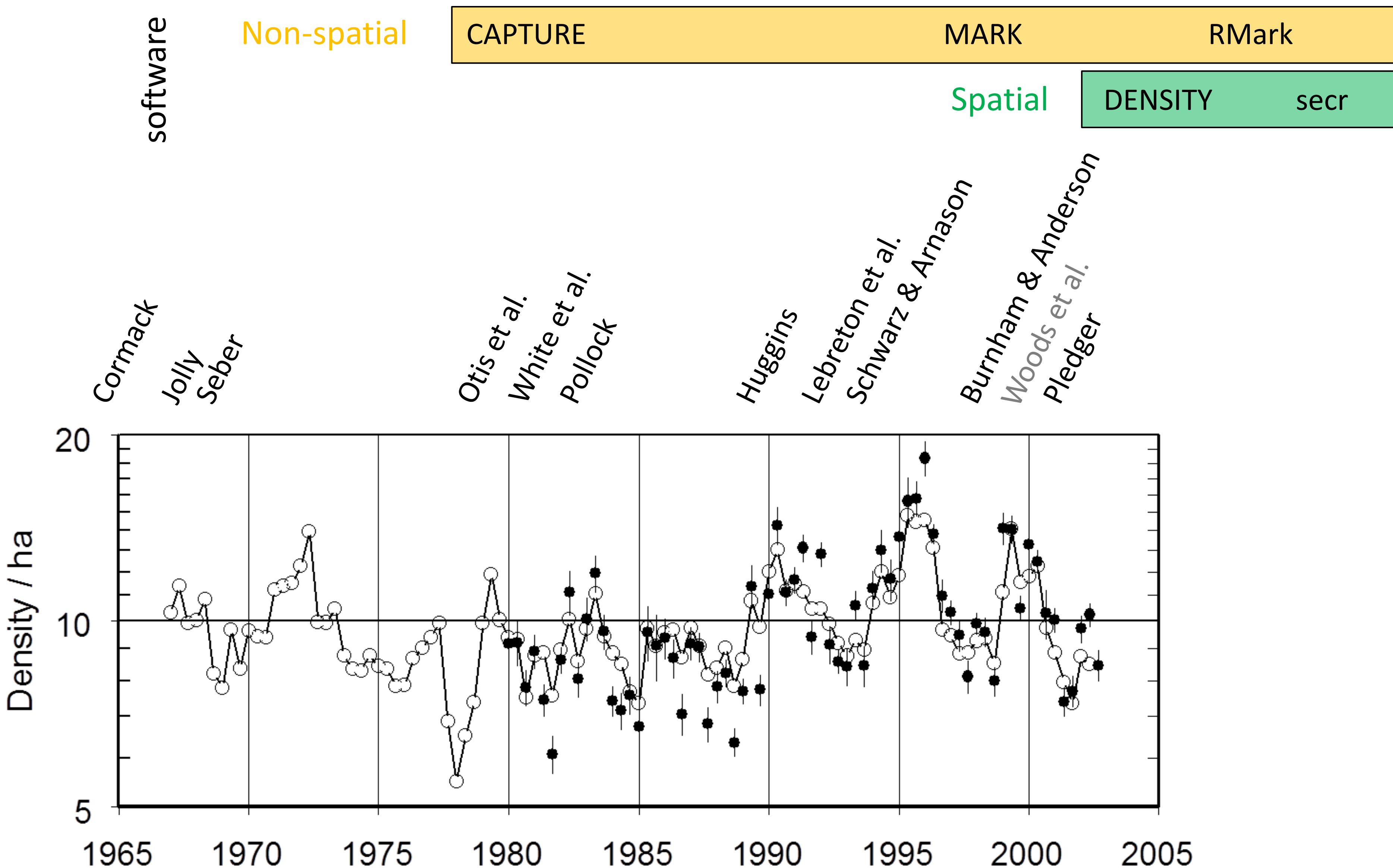


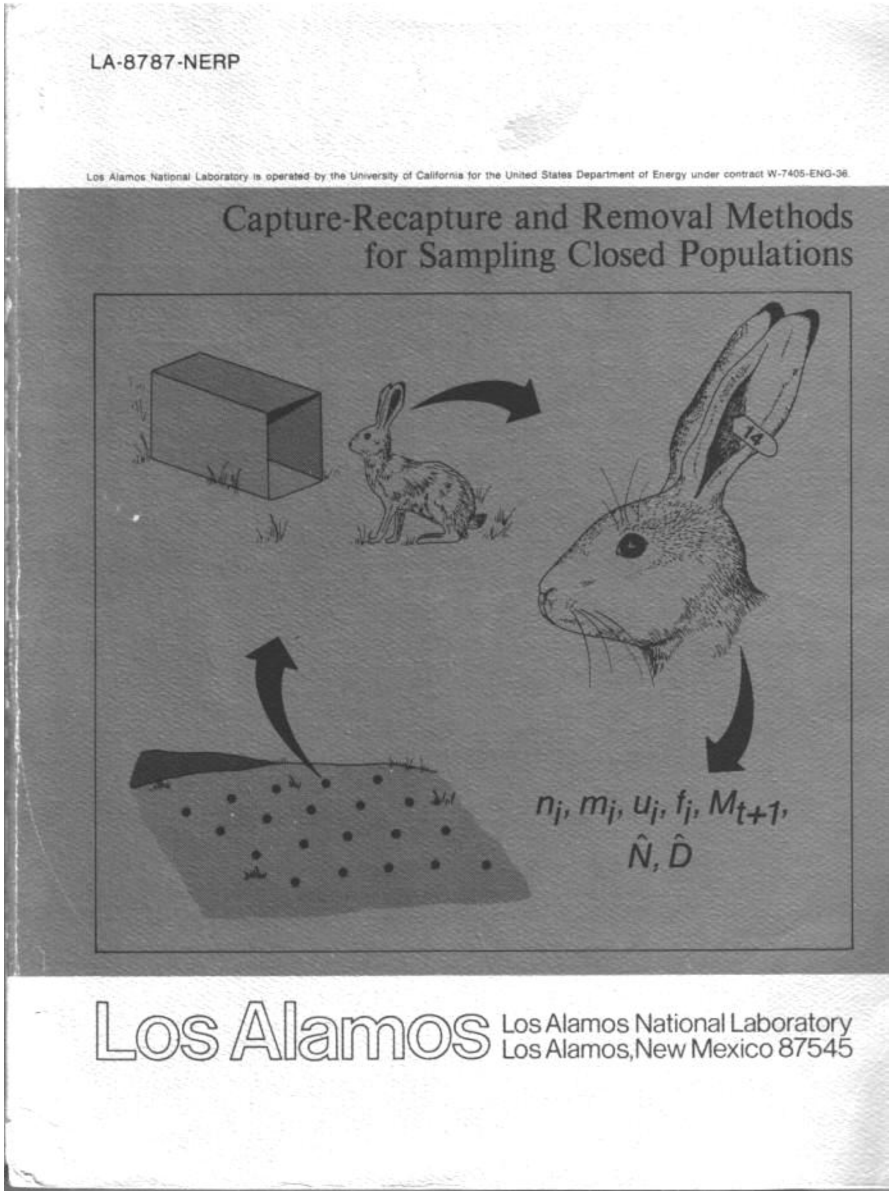




Efford MG, Cowan PE 2004. Long-term population trend of *Trichosurus vulpecula* in the Orongorongo Valley, New Zealand. In: The Biology of Australian Possums and Gliders. Edited by RL Goldingay and SM Jackson. Surrey Beatty & Sons, Chipping Norton. Pp. 471–483.

A partial history of capture-recapture…





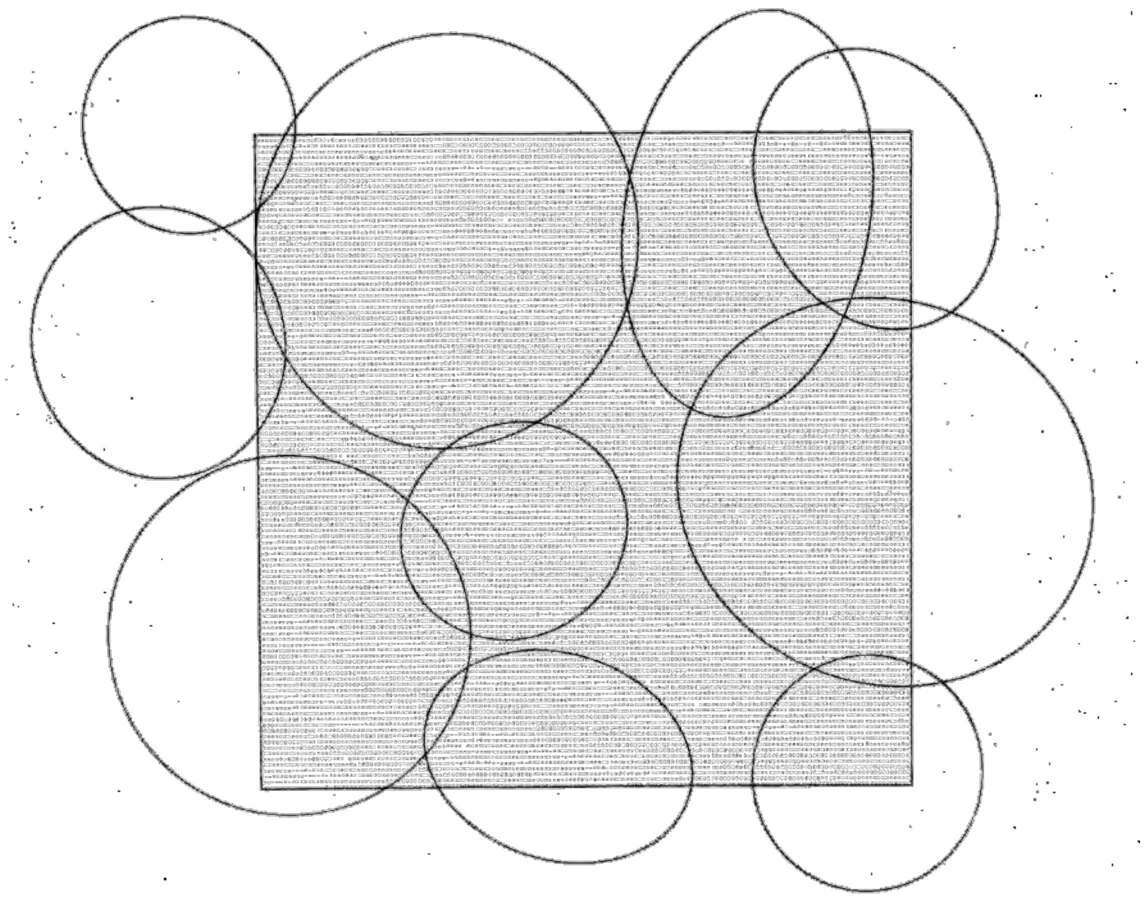
White et al. (1982)

ANIMAL
ABUNDANCE

POPULATION
SIZE N

DENSITY D

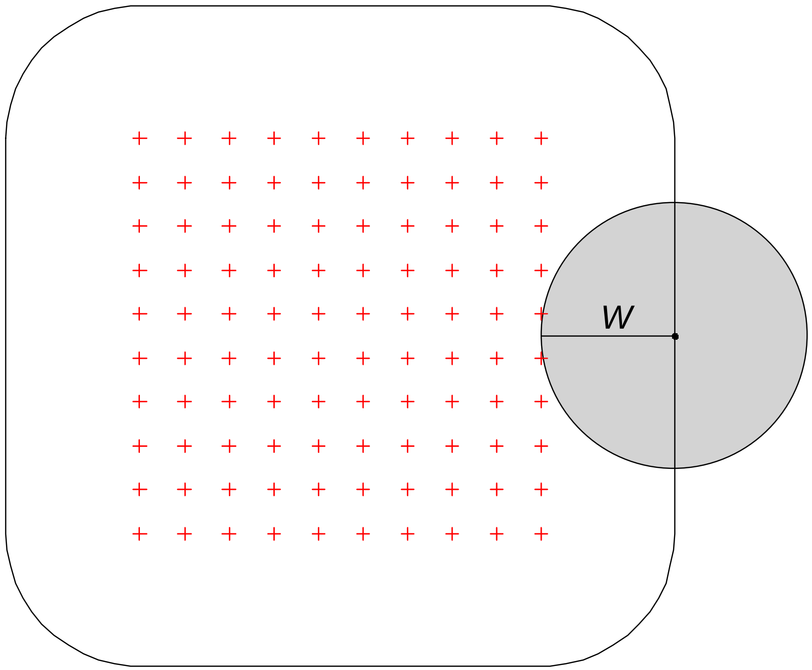
White et al. 1982



"Fig. 5.3. Because almost all of the home ranges (ellipses) include some area outside the trapping grid (shaded area), the grid's effective area is much larger than its physical area. At best, a very poor density estimate would be achieved under these circumstances."

Bias of naive estimator: about +100%

Density estimation : Boundary strip method



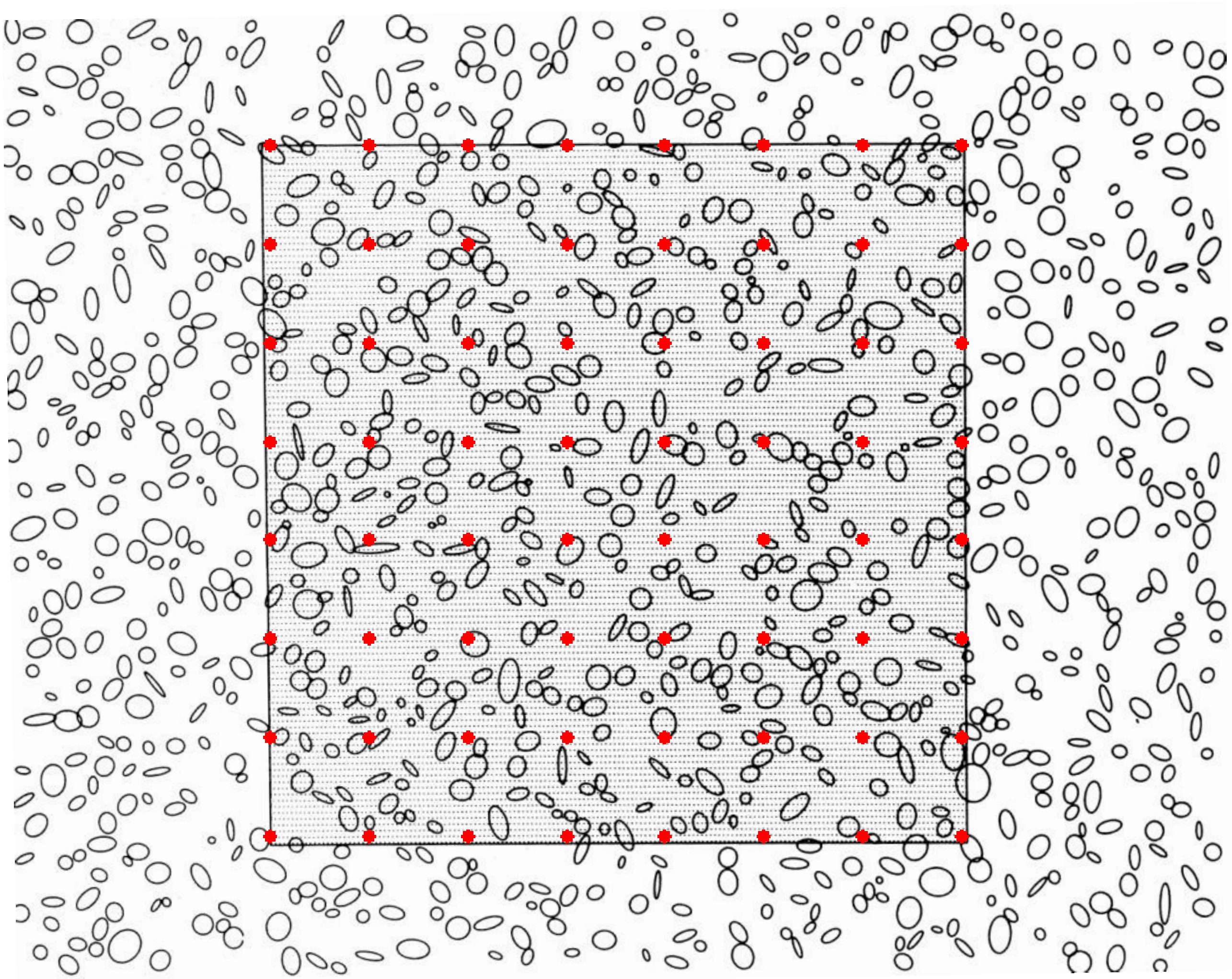
Conventional estimator

$$\hat{D} = \hat{N} / A_W$$

"Effective trapping area" A_W

What is W? W = MMDM/2? W = MMDM?
Come to think of it, what is N?

White et al. 1982 Ideal design: large grids and small ranges



Observe: 95% of home ranges do not include a trap

Bias of naive estimator: about -95%

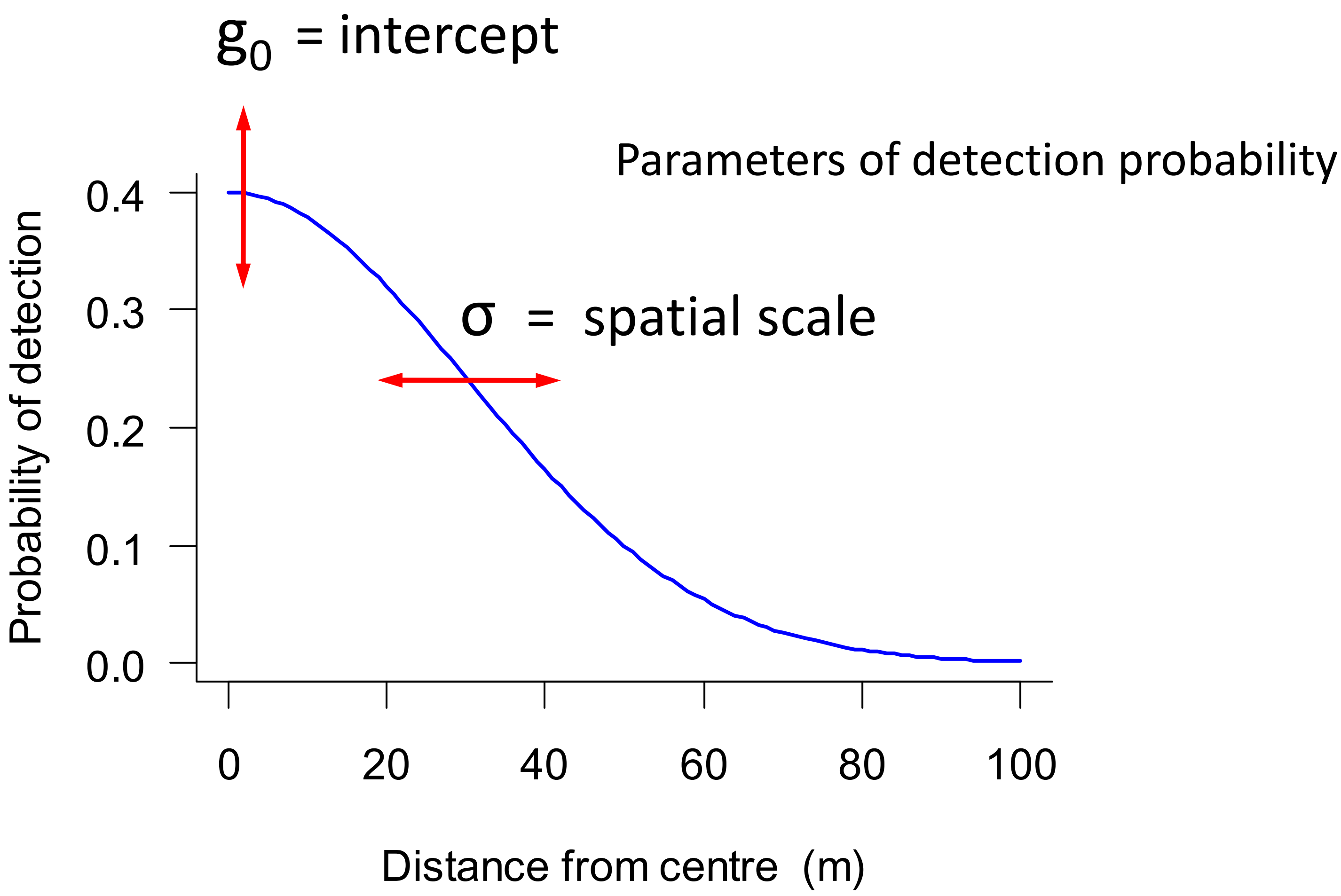
The movement paradox

1. Movement blurs the definition of the sampled population

2. Passive detectors rely on movement

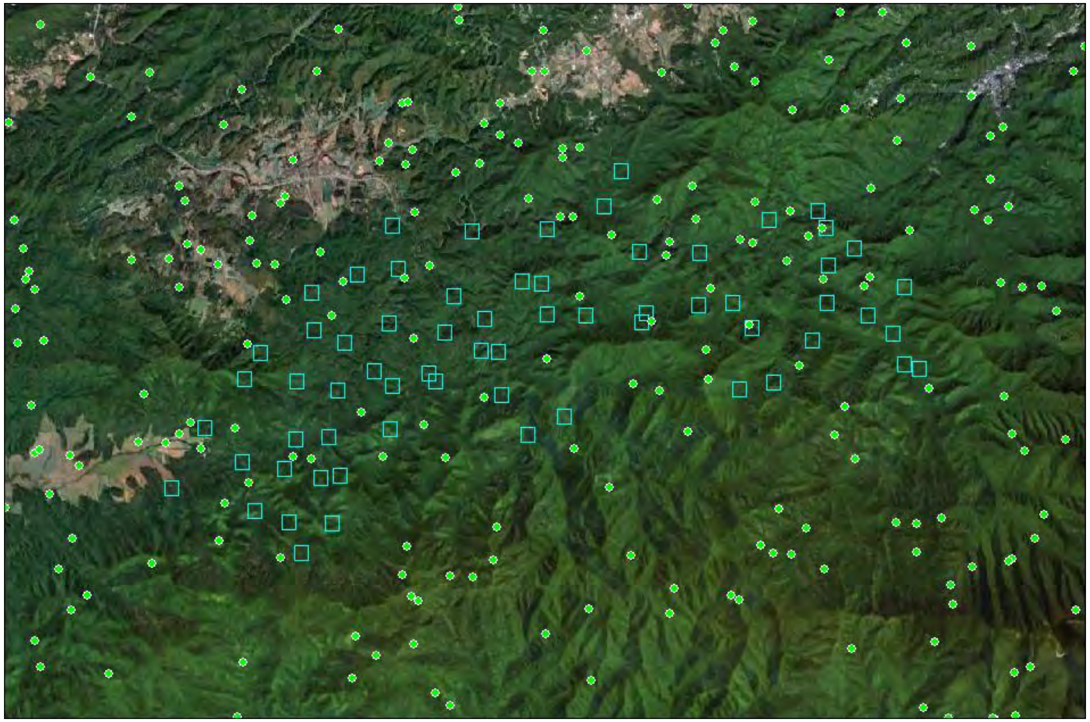
SECR solution: live with movement by including it in model

Including movement in the model: distance-dependent detection



Conventional parameters	*N*, *p*
SECR parameters	D, g_0, σ

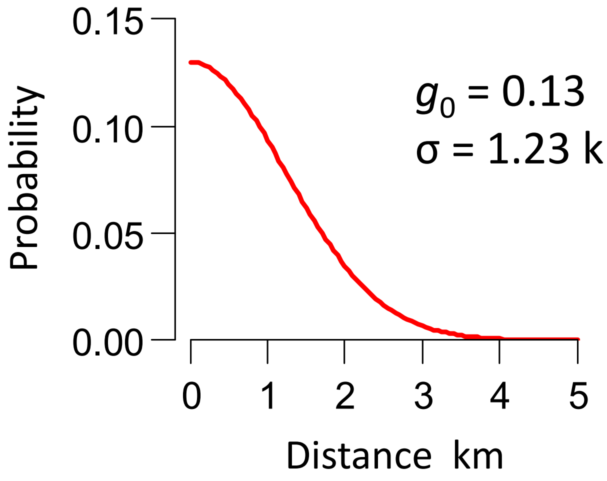
Green dots : Poisson distribution with fitted density



GSM black bears: data of Jared Laufenberg , Frank van Manen, and Joe Clark

Fitted model

Density 0.32 / km² (0.24 – 0.42 / km²)



Summary: What is SECR good for?

	Software* DENSITY	secr
Estimating population density without edge effects	●	●
Testing survey designs	●	●
Estimating population size in a defined region		●
Relating density to habitat, time etc.		●

All difficult or impossible with non-spatial methods

* see Appendix of secr-overview.pdf for detailed comparison

4. What are the limitations of SECR?

- Computationally intensive
- It's still capture-recapture
  - Good to have plenty of data
  - Poor model selection may or may not lead to bias
  - Too many models to choose from
- Under development
  - Overdispersion estimation and goodness-of-fit tests
  - Semi-parametric surfaces
  - Open-population and mixed-data methods
  - Documentation of robustness (transients and elongated home ranges – effects not usually severe)

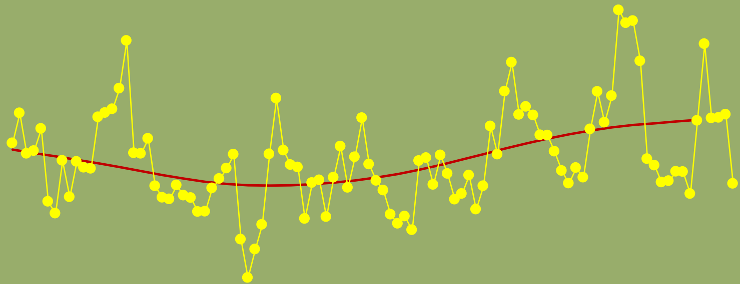
Murray Efford
Western Black Bear Workshop
Coeur d'Alene ID May 2012

Morning session

1. Introduction

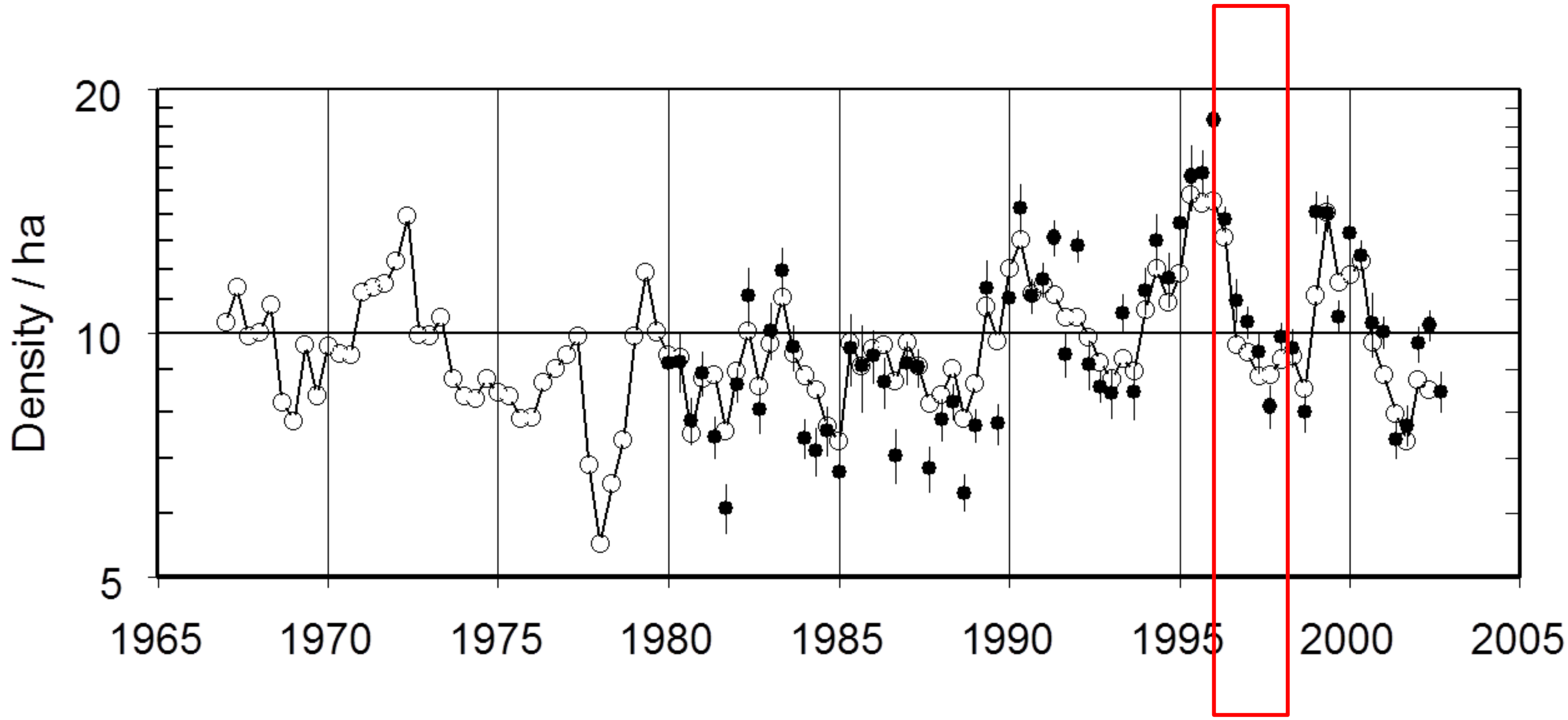
2. DENSITY software
    - Interface
    - Data
    - Conventional analyses
    - Simple spatial analysis: GSM black bears

3. Key concepts
    - Detector types
    - Buffers, habitat masks, and the 'region of integration'
    - Maximum likelihood



Brushtail possum (*Trichosurus vulpecula*)
Capture-recapture on ca. 14-ha grid
Orongorongo Valley, New Zealand

Sessions 49-54
1996, 1997



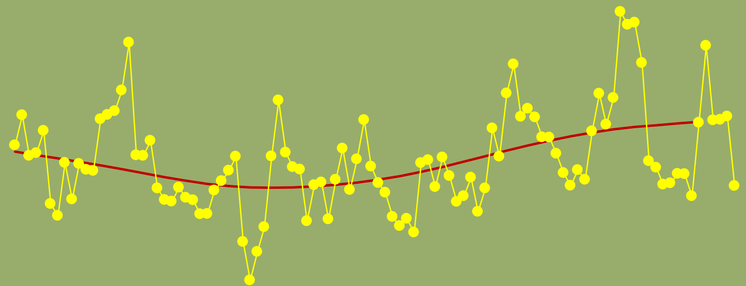
ovtrap.txt, ovcapt.txt

Morning session

1. Introduction

2. DENSITY software
   - Interface
   - Data
   - Conventional analyses
   - Simple spatial analysis: GSM black bears

3. Key concepts
   - Detector types
   - Buffers, habitat masks, and the 'region of integration'
   - Maximum likelihood etc.



Detector types

- SECR models the probability of detection <u>at each detector</u>
- Different types of detector require slightly different models


Photograph Bruce Warburton

Detector types

Effect of capture event on :

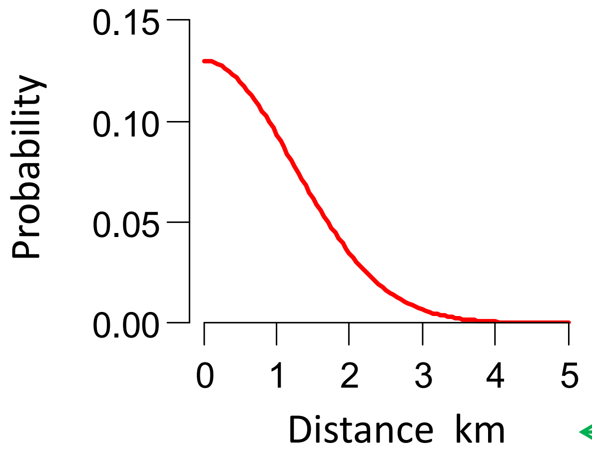
	Animal	Trap
Single-catch trap[1]	trapped	full
Multi-catch trap[2] pitfall, mist net	trapped	available
Proximity detector[3] camera, hair snag	free	available

1. No likelihood available
2. Competing risk (hazard) likelihood Borchers and Efford 2008
3. Detectors independent

Detector types

- SECR models each detector
- Different types of detector require slightly different models

Buffers and habitat masks



Distance from detectors to home-range centers

PROBLEM: we don't know where the centers are

SOLUTION: consider all <u>possible locations</u>*, weighting by probability

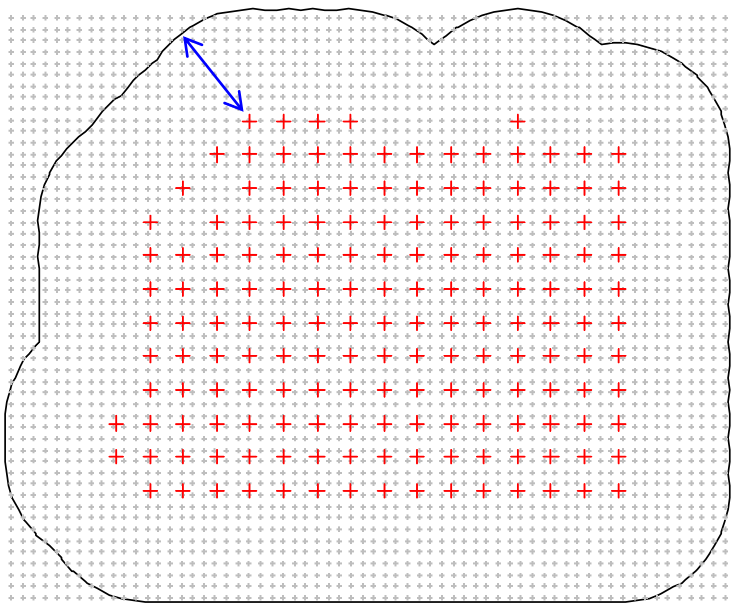
(integrated likelihood OR MCMC)

* possible locations = habitat mask = region of integration = state space

‘Possible locations’ for centers of detected animals

1. All points within an arbitrary ‘buffer’ radius of detectors…

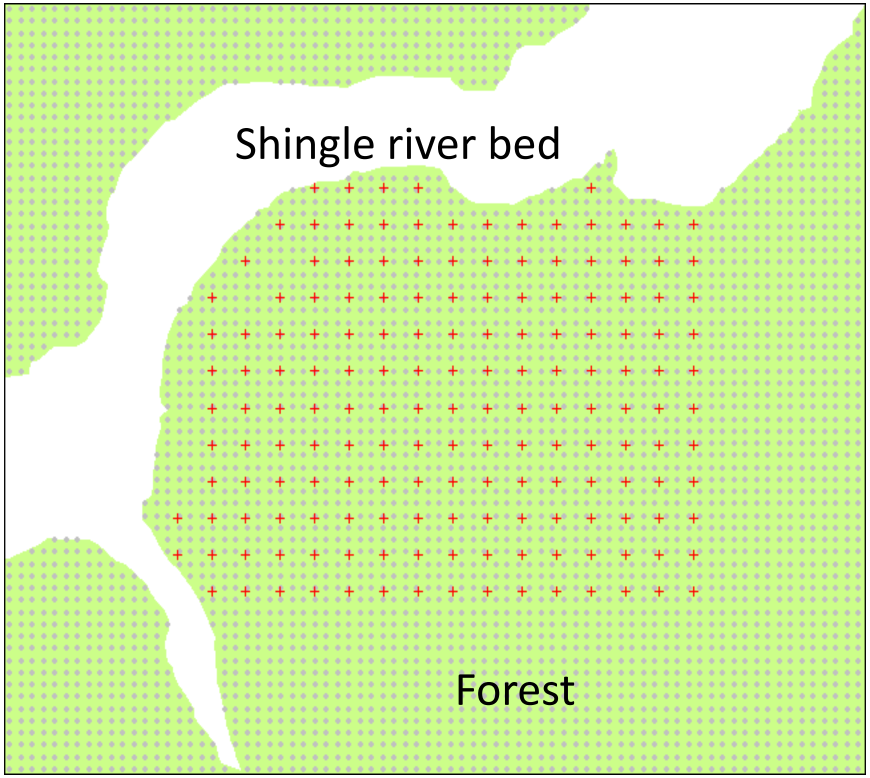
(where buffer is greater than any likely movement distances)



If the buffer is too narrow then bias may result – there are post hoc methods to recognise this.

'Possible locations' for centers of detected animals cont'd

2. As before, but with other biologically justified constraints, e.g.



Maximum-likelihood estimation

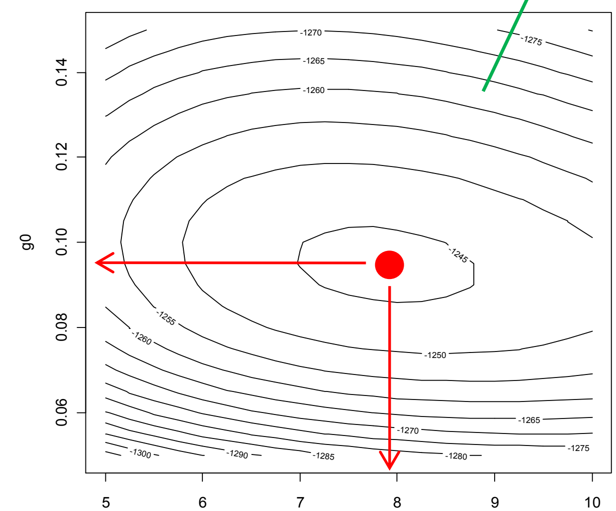
- Likelihood can be calculated from <u>data</u> for given <u>parameter values</u>

- Maximum likelihood corresponds to ‘best’ parameter estimates

- Use numerical (computer) methods* to find maximum, given some starting values

likelihood surface

g_0



Density

g_0 and Density are model parameters

* alternative algorithms: Newton-Raphson, Nelder-Mead, BFGS

Two ways of fitting SECR model

1. Maximize <u>full likelihood</u> $\longrightarrow$ $\widehat{D},\ \hat{g}_0, \hat{\sigma}$

2. Maximize <u>conditional likelihood</u> $\longrightarrow$ $\hat{g}_0, \hat{\sigma}$ just the detection parameters

$$\hat{a} = \int_R p.(x; \hat{g}_0, \hat{\sigma})dx$$

'effective sampling area' in sense of Borchers & Efford 2008

$$\widehat{D} = n/\hat{a}$$

Horvitz-Thompson-like estimate cf Huggins 1989

number of unique individuals detected

Full vs conditional likelihood

Full
- Density is a model parameter
- Allows modelling of density between sessions or vs habitat (secr)
- Allows profile-likelihood confidence interval on density
- Individual covariates prohibited except as ‘groups’ or ‘sessions’

Conditional
- Density is a derived variable, not a model parameter
- Allows any individual covariate, continuous or categorical
- Allows spatial variance to be estimated empirically
- Simpler likelihood
- Sometimes faster

Two forms, but nearly identical estimates of density: choose to suit your problem

The ‘Distribution’ setting: Two ways to conceive target population

	1.	2.
N*	Poisson	Fixed
n	Poisson	Binomial
	Expected	Realized
	‘Cookie-cutter’ segment of extensive pattern	Specific realization of spatial process

* population in region of integration

Excludes ‘process’ variance, so SE smaller

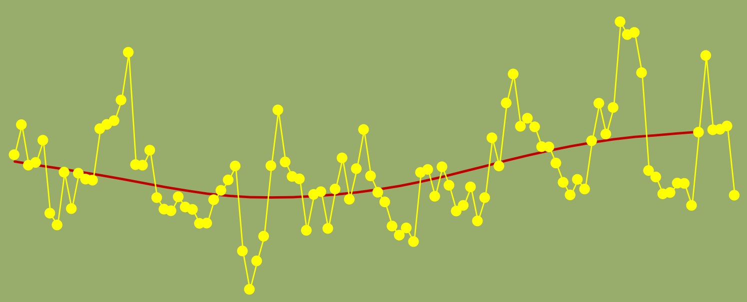
Connects with scope of inference – see ‘study design’

Afternoon session

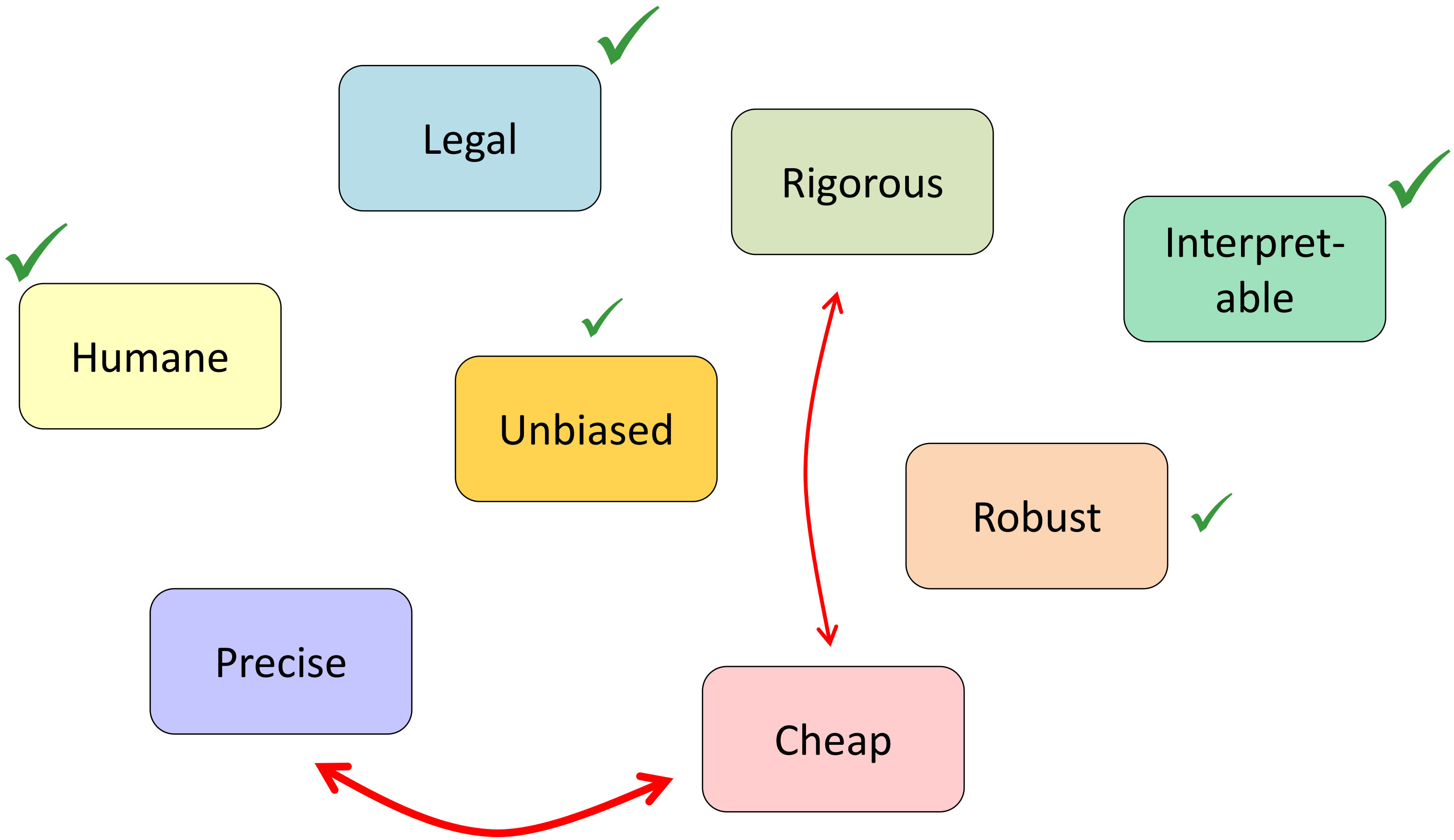
4. Study design
    - Design goals
    - Spatial representativeness
    - Simulation
    - Composite designs
    - Rules of thumb

5. R package 'secr'

6. Miscellany



Design goals for capture-recapture monitoring



✓SECR compatible

Trade-offs

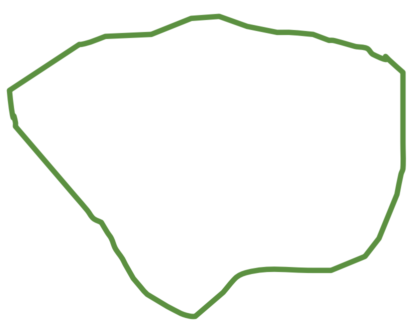
Rigorous?

A rigorous monitoring design has

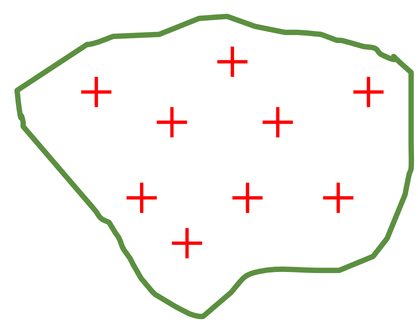
- well-defined state variable(s)
- credible estimates of precision
- explicit scope of inference

SECR generally delivers these

defined region of interest*



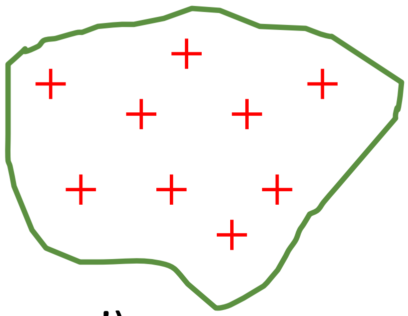
spatially representative sampling



* may be much larger than SECR region of integration

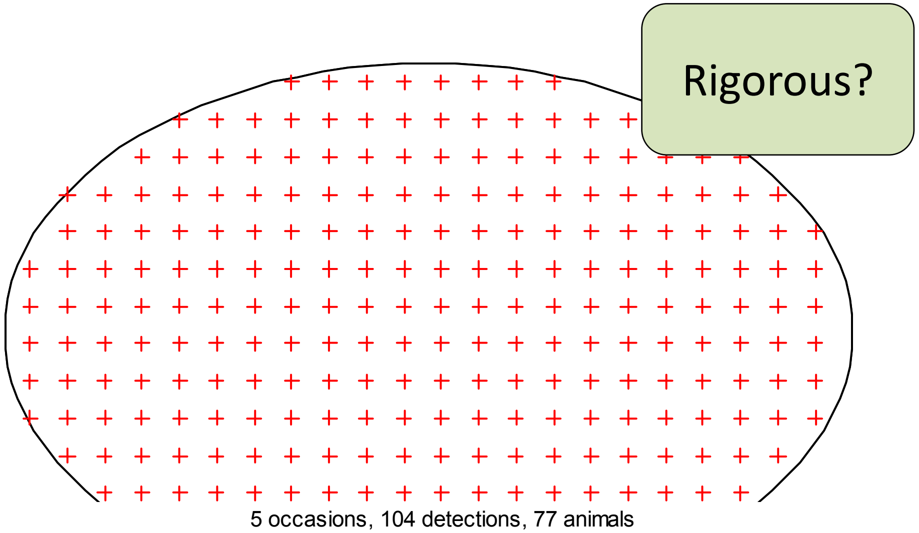
Rigorous?

Principles of spatially representative sampling

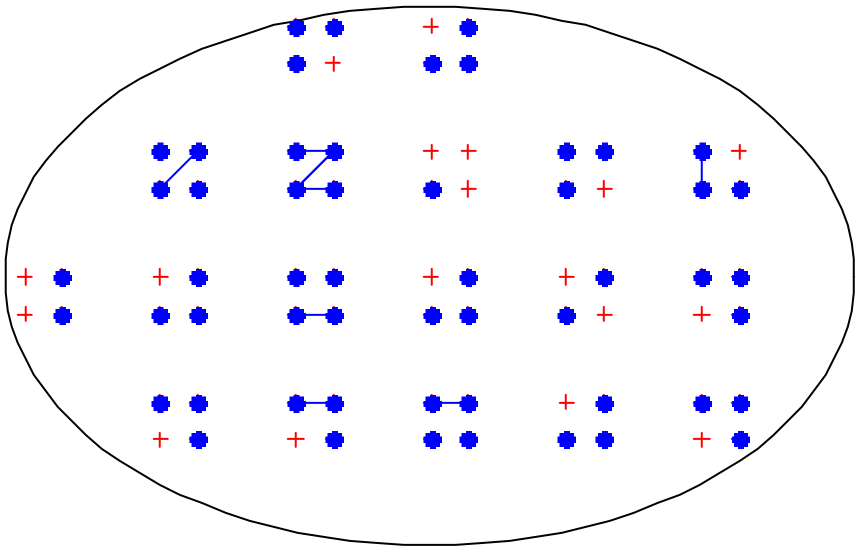


- Probability-based sampling options:
  - Simple random
  - Systematic with random origin
  - GRTS (Stevens & Olsen 2004; package 'spsurvey')
- Identify & exclude inaccessible areas
- Stratify to reduce cost
- Refer sampling literature and Distance books

Options for spatially representative sampling of large regions



A. Continuous grid



B. Clusters (mini-grids)

Recaptures mostly within clusters

Rigorous?

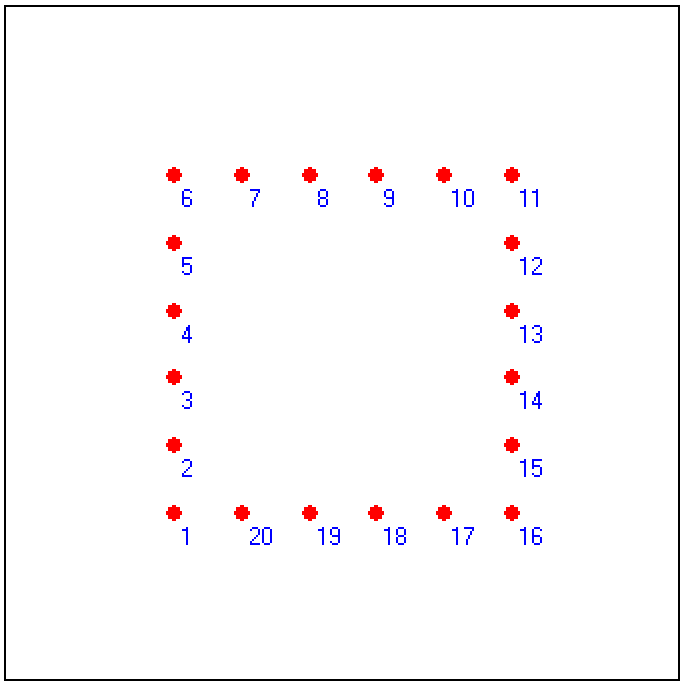
<u>Clustering of detectors</u> can be a good compromise, allowing researcher to:

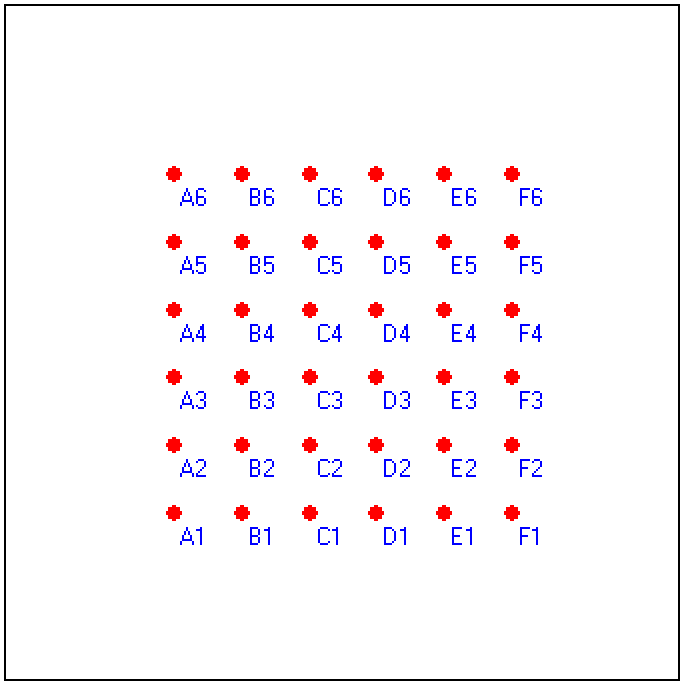
- Sample a region rigorously by placing clusters according to a probability-based design
- Maintain healthy distribution of potential recapture distances within clusters

Increase number of clusters to increase sampling effort and precision

Cheap!

Some cluster designs (e.g. hollow grids) are attractive for logistical reasons: fast to lay out and efficient to operate





…but linear, road-side surveys require careful justification

Precise?

Precision means

- Small relative SE (= 'CV')
- Short confidence/credible intervals
- High power to detect change

Precise?

Components of variance

$$\text{var}(\hat{D}) \approx \hat{D}^2[\text{CV}^2(n) + \text{CV}^2(a(\hat{\theta}))]$$

encounter rate uncertainty

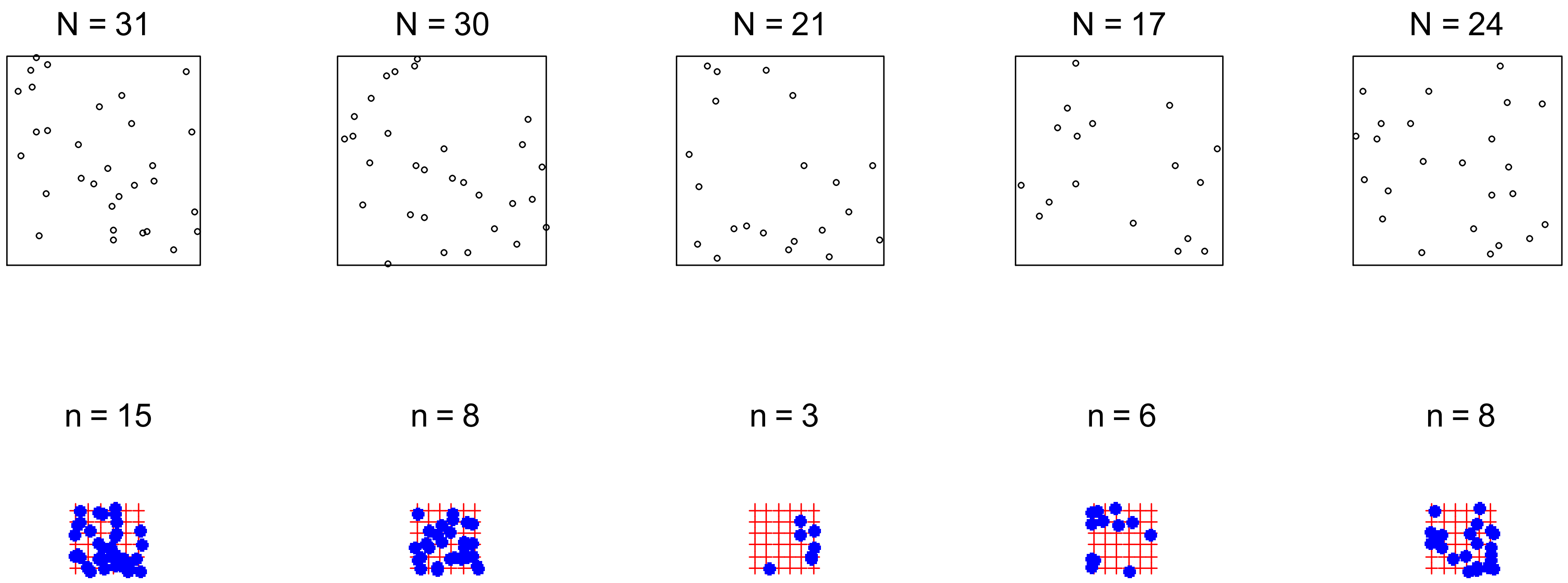
detection function uncertainty

Which is dominant?

Precise?

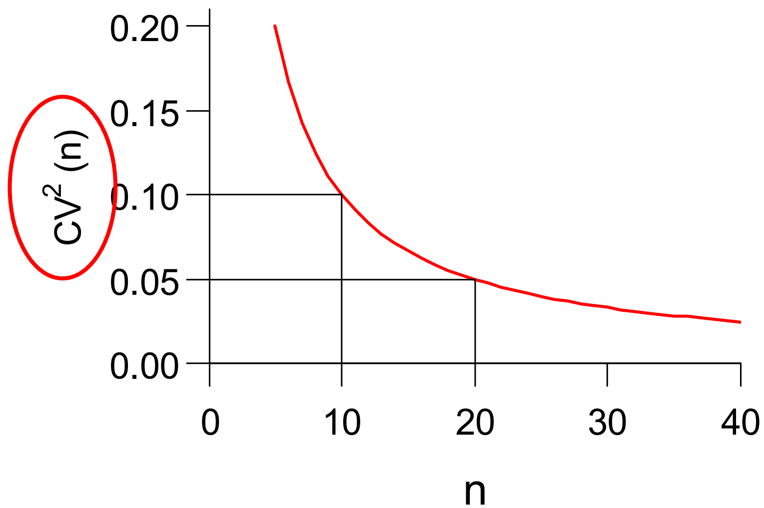
encounter rate uncertainty = chance variation in the number of animals observed

Example : uniform global density = 3 / ha, but samples vary



Precise?

The Poisson floor: 1/*n*

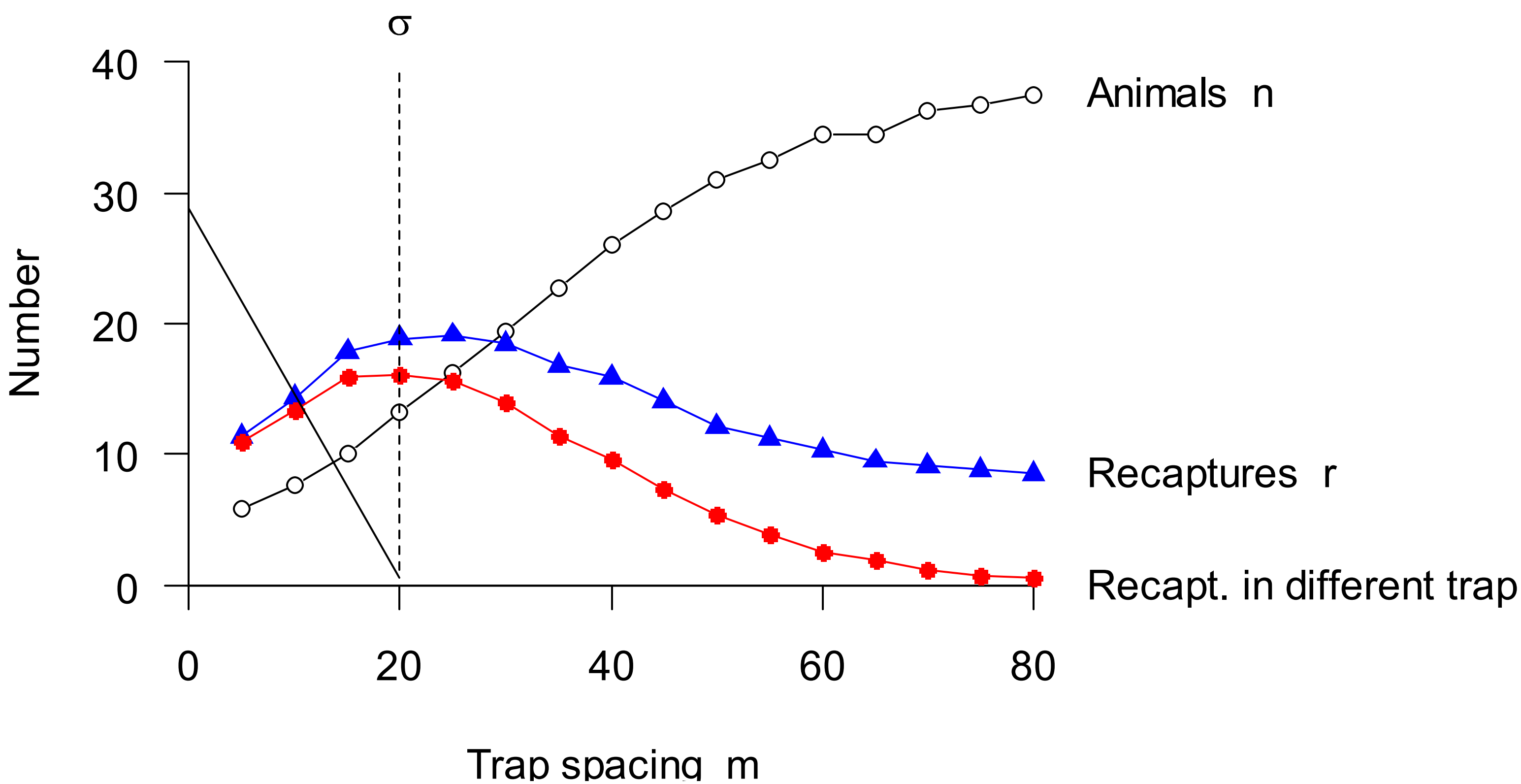


$$\mathrm{var}(\hat{D}) \approx \hat{D}^2[\mathrm{CV}^2(n) + \mathrm{CV}^2(a(\hat{\theta}))]$$

Estimates of sparse populations with low detection rates are imprecise, <u>regardless</u> of how well detection function is estimated

How trap spacing affects precision

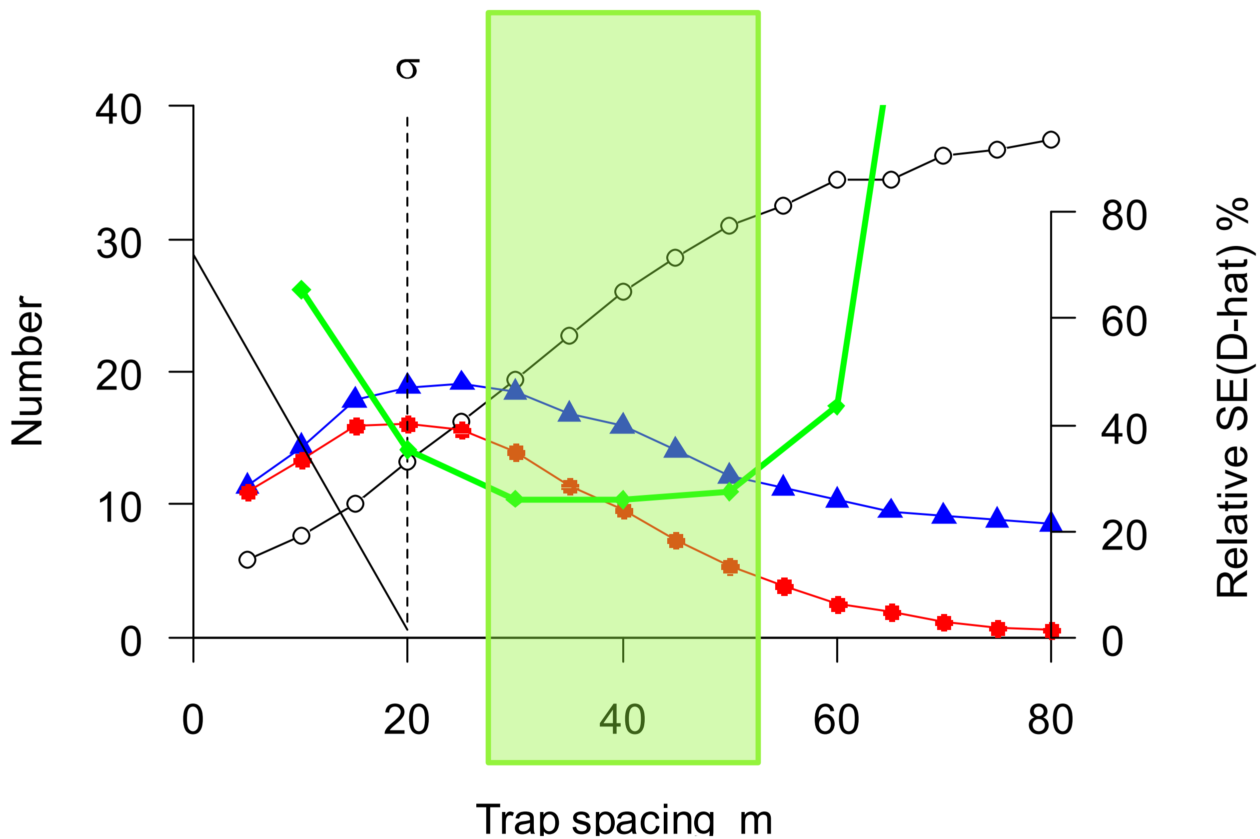
Precise?



Widely spaced traps yield large *n*, but small *r*.

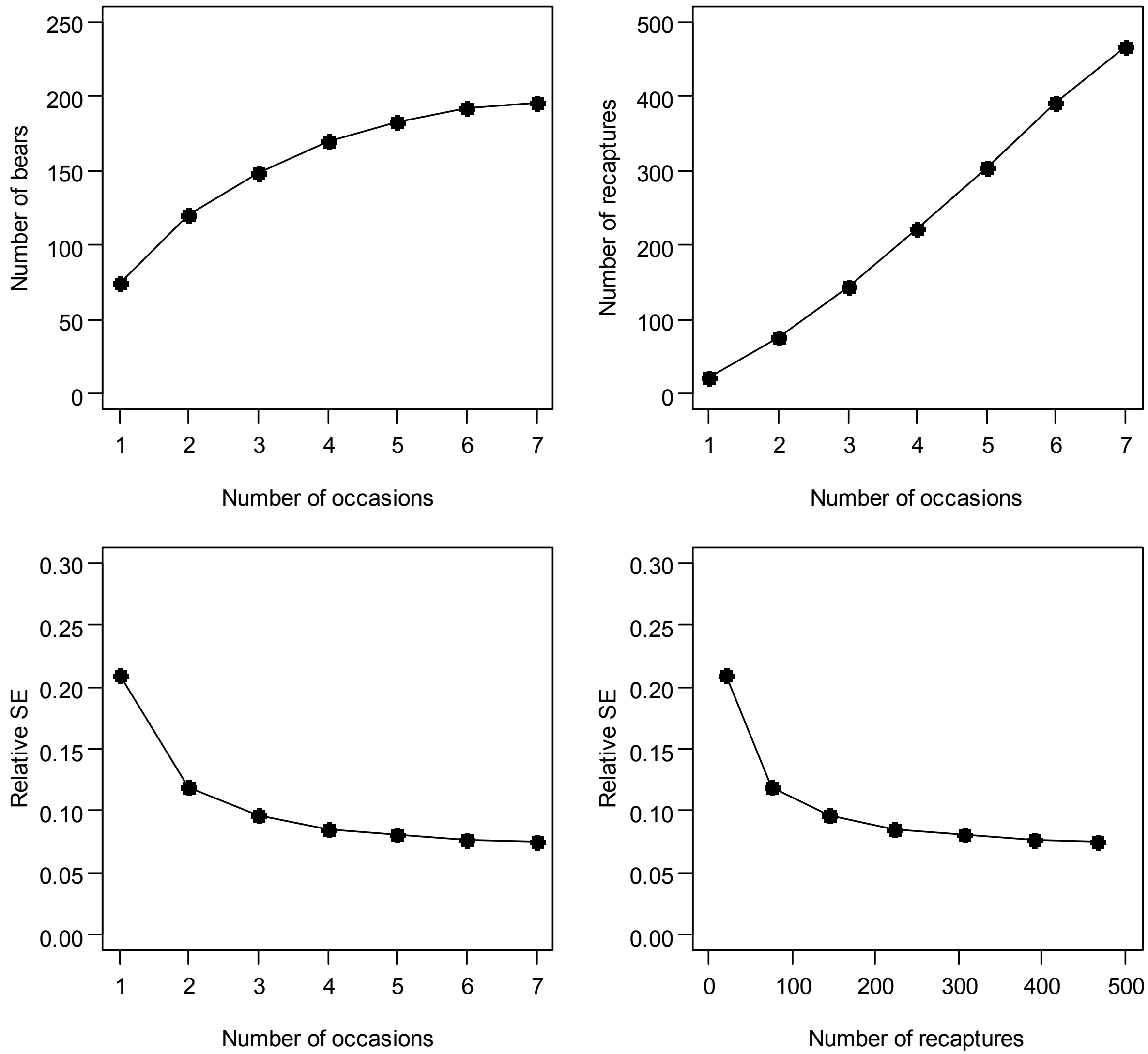
Precise?

Precision is best at intermediate spacing
(here 1.5 σ – 2.5 σ)



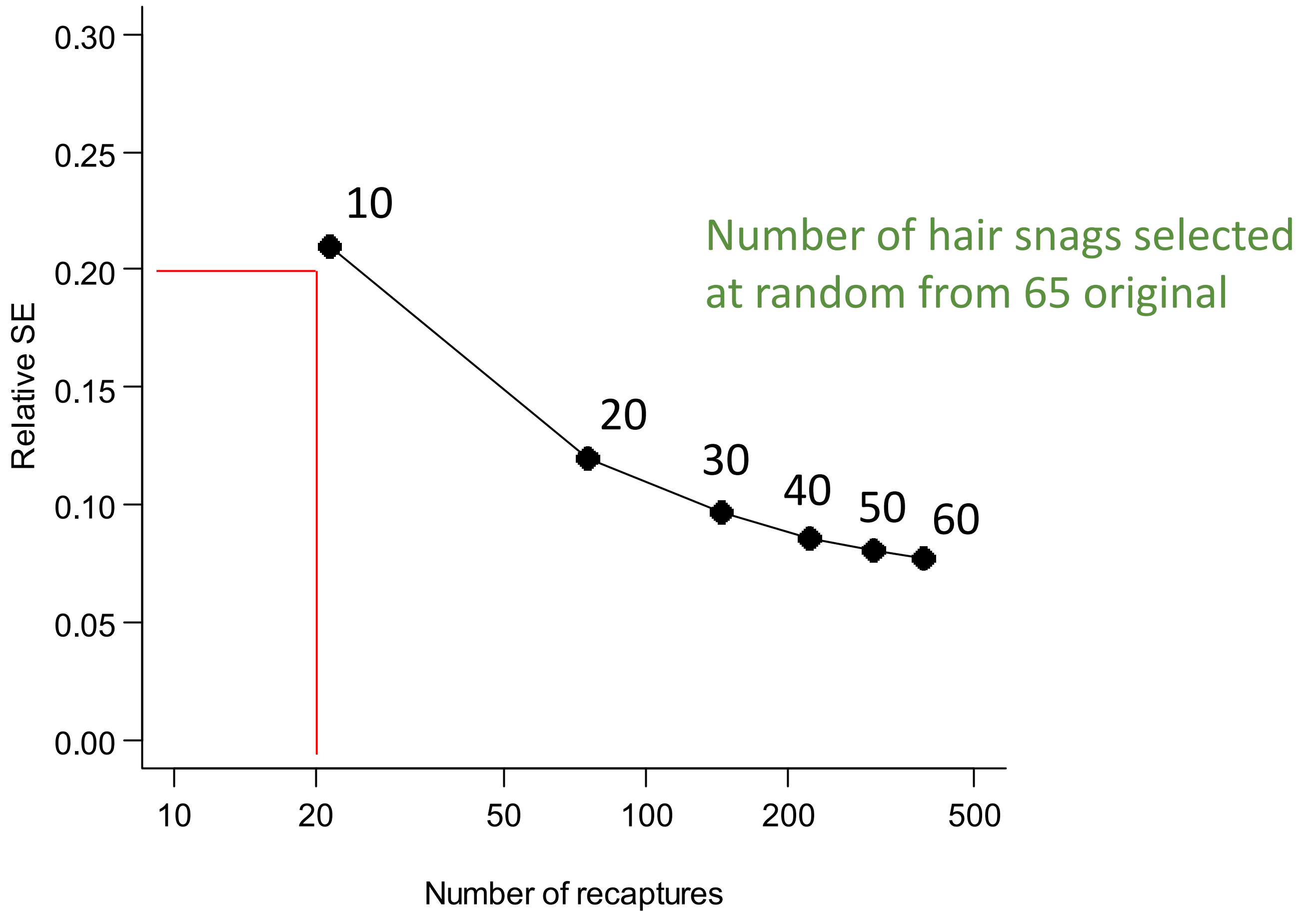
GSM bear simulations D = 0.8 / km^2, g_0 = 0.13, σ = 1.5 km

Precise?



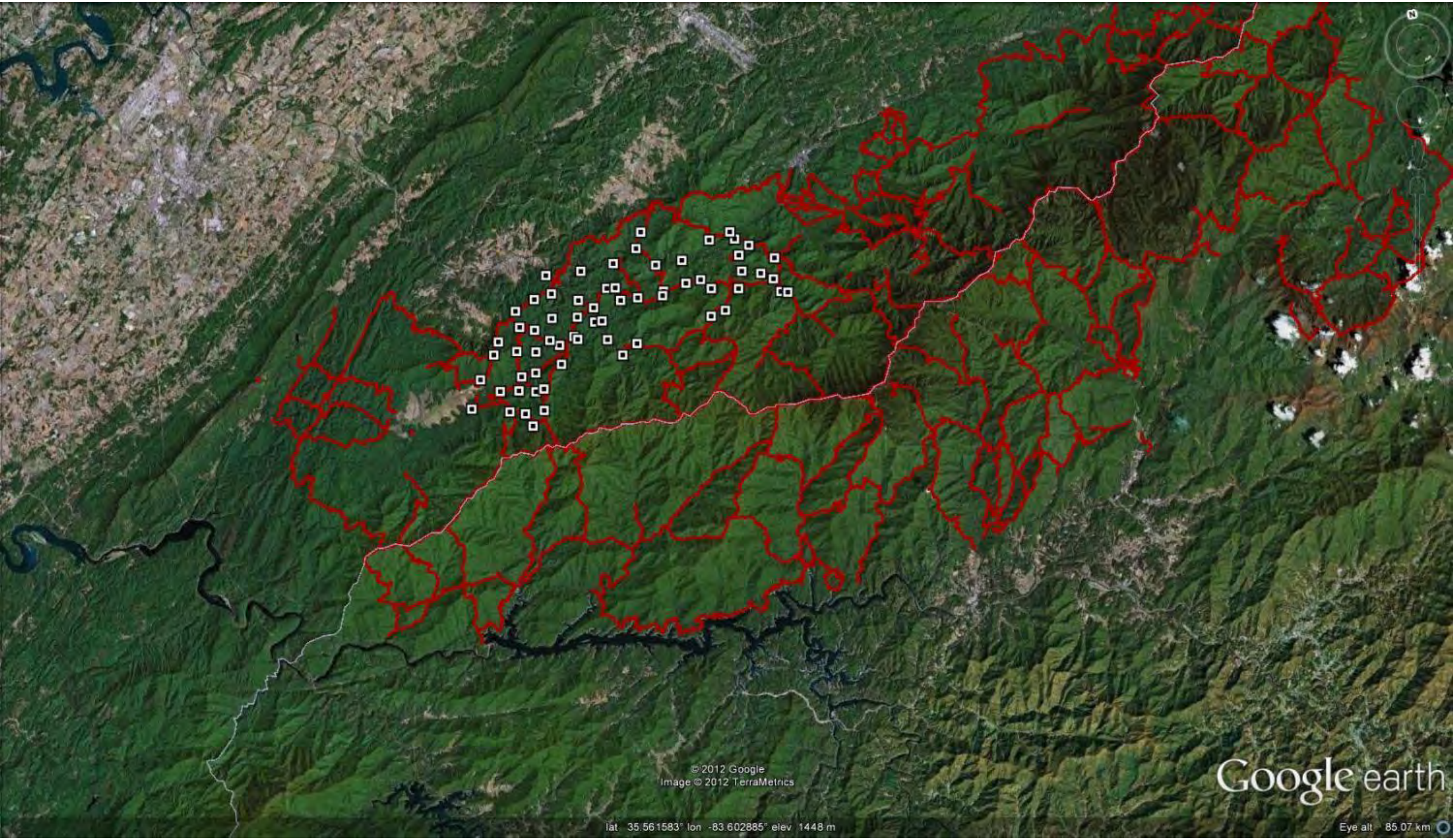
GSM bear simulations $D = 0.8 / km^2$, $g_0 = 0.13$, $\sigma = 1.5$ km - varying number of detectors

Precise?



20:20 rule of thumb

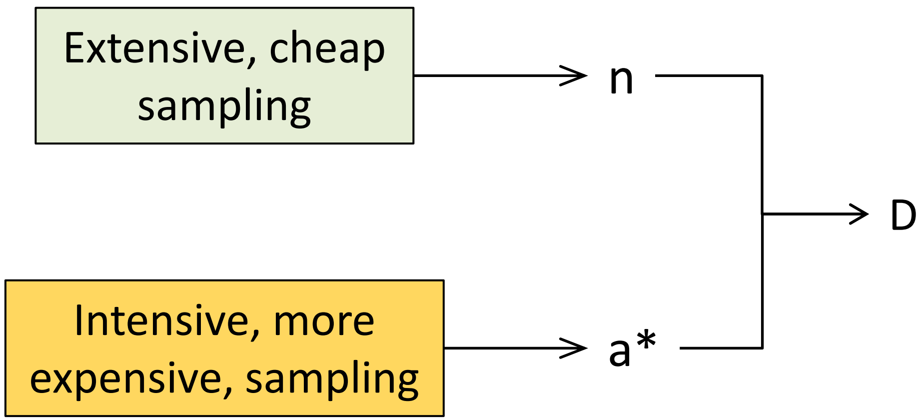
GSM revisited – cover larger region of interest with same number of detectors?



Precise?

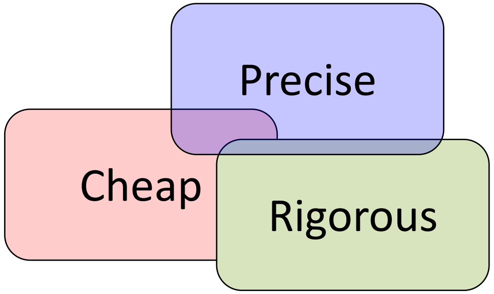
Rigorous?

Composite designs



* a = effective sampling area = integrated detection probability

Rigorous selection of intensive sites is essential



Study design summary

Define region of interest

Spatially representative sample

Cluster detectors for flexibility

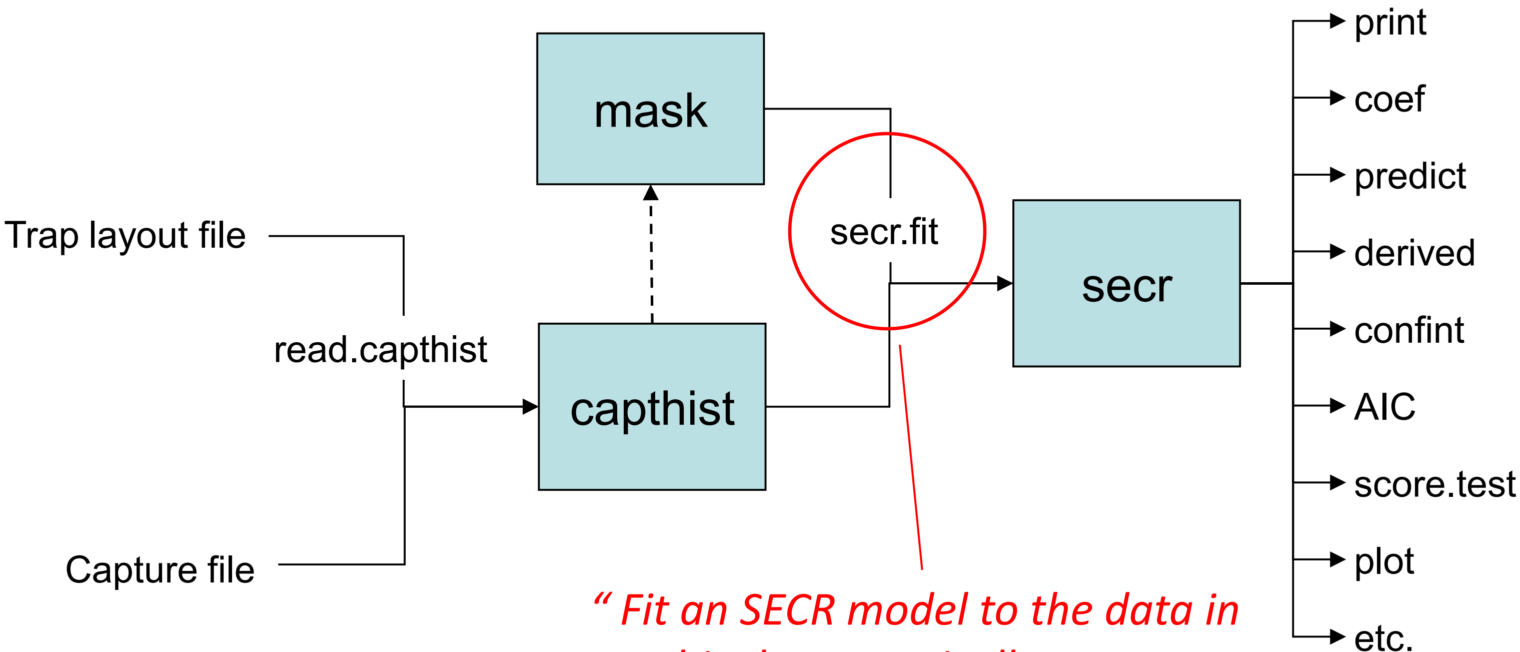
Test by simulation

Rules of thumb - 2σ spacing, >20 recaptures

Consider composite designs

	DENSITY	secr
Graphic interface	✓	
Simulation manager	✓	
Windows OS	✓	✓
Other OS		✓
Advanced models		✓
Scripts		✓
	32-bit	32-bit or 64-bit (faster, more memory)

Mastering secr.fit()



" Fit an SECR model to the data in capthist by numerically maximising the likelihood; return an object of class 'secr' "

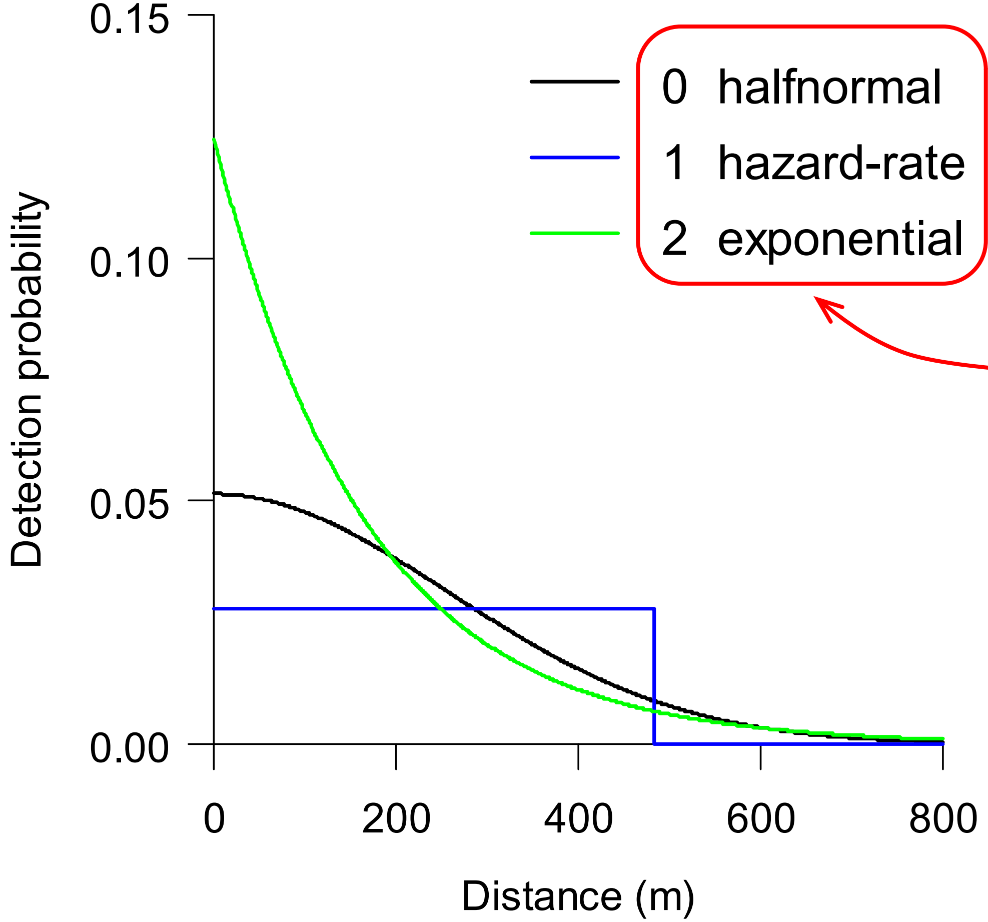
The simplest possible analysis

```
library (secr)
setwd (system.file("extdata", package = "secr"))
stoatCH <- read.capthist ("stoatcapt.txt",
          "stoattrap.txt“, detector = "proximity")

secr.fit (stoatCH, buffer = 1000)
```

Implied (default) arguments –

CL = F	maximise full likelihood
detectfn = 0	halfnormal detection function
mask	automatic (buffer = 1000 m)
start	automatic initial values for parameters
model = list(D~1, g0~1, sigma~1)	constant model

Detection functions



$$g(d) = g_0 \exp(-d^2/(2\sigma^2))$$

$$g(d) = g_0 (1 - \exp(-(d/\sigma)^{-z}))$$

$$g(d) = g_0 \exp(-d/\sigma)$$

See ?detectfn for numeric codes

Fitted to stoatCH

Comparing density estimates

```
data(stoatDNA)
options(digits=3)

collate(stoat.model.HZ, stoat.model.HN, stoat.model.EX,
      realnames='D', perm=c(2,3,4,1))

, , D, session=MatakitakiStoats

              estimate SE.estimate    lcl    ucl
stoat.model.HZ   0.0234     0.00682 0.0134 0.0410
stoat.model.HN   0.0224     0.00654 0.0128 0.0392
stoat.model.EX   0.0223     0.00661 0.0126 0.0394
```

Negligible difference between hazard-rate, halfnormal and exponential detection functions

The model specification

- One formula for each ‘real’ parameter –

model = list($\underbrace{\text{D~1}}_{\boxed{\text{density model}}}$, $\underbrace{\text{g0~1, sigma~1}}_{\boxed{\text{detection model}}}$)

later

Formulae use R notation for linear models – see help(formula)	*For example*	
	~1	constant
	~x	linear fn of x
	~x+y	additive linear fn of x and y

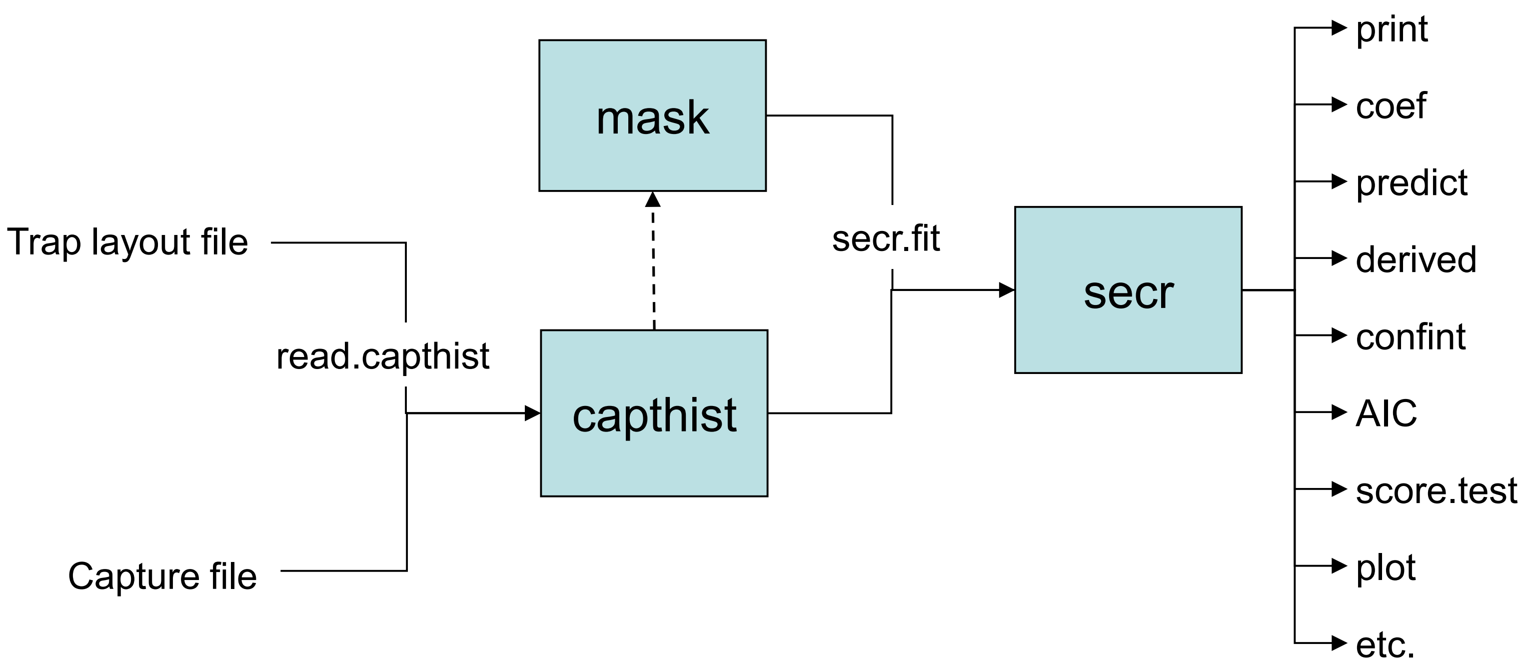
Possible terms in the detection model

only with full likelihood

Term	Description	Notes
g	group factor	interaction of the capthist individual covariates listed in argument 'groups'
t	time factor	one level for each occasion
T	time trend	linear trend over occasions on link scale
b, bk	learned response	step change in real parameter after first detection of animal (bk site-specific)
B, Bk	transient response	real parameter depends on detection at previous occasion (Markovian response)
session	session factor	one level for each session
h2	2-class mixture	finite mixture model with 2 latent classes

These are available automatically; others may be supplied as covariates

secr.fit() returns an ‘secr’ object



Do not look directly at an secr object (unless you really have to)!

An secr object is a list with 26 components

```
data(stoatDNA)
names(stoat.model.HN)

 [1] "call"        "capthist"    "mask"        "detectfn"
 [5] "CL"          "timecov"     "sessioncov"  "groups"
 [9] "dframe"      "design"      "design0"     "start"
[13] "link"        "fixed"       "parindx"     "model"
[17] "details"     "vars"        "betanames"   "realnames"
[21] "fit"         "beta.vcv"    "D"           "version"
[25] "starttime"   "proctime"
```

print.secr makes a readable summary

call

```
secr.fit( capthist = stoatCH, buffer = 1000, detectfn = 0 )
secr 1.4.0, 16:35:36 03 May 2010
```

data

```
Detector type     multi
Detector number   94
Average spacing   250 m
x-range           -1500 1500 m
y-range           -1500 1500 m
N animals       :  20
N detections    :  30
N occasions     :  7
Mask area       :  2500 ha
```

model

```
Model           :  D~1 g0~1 sigma~1
Fixed (real)    :  none
Detection fn    :  halfnormal
Distribution    :  poisson
N parameters    :  3
Log likelihood  :  -144.0016
AIC             :  294.0033
AICc            :  295.5033
```

Estimated density
animals / hectare

coefficients
(on link scale)

```
Beta parameters (coefficients)
           beta    SE.beta       lcl       ucl
D     -3.800341 0.2865730 -4.362014 -3.238668
g0    -2.913927 0.4445352 -3.785200 -2.042654
sigma  5.552586 0.1721433  5.215191  5.889981

Variance-covariance matrix of beta parameters
                 D          g0        sigma
D      0.082124067 -0.04108776 -0.007142058
g0    -0.041087764  0.19761153 -0.054651267
sigma -0.007142058 -0.05465127  0.029633332
```

'real' parameters

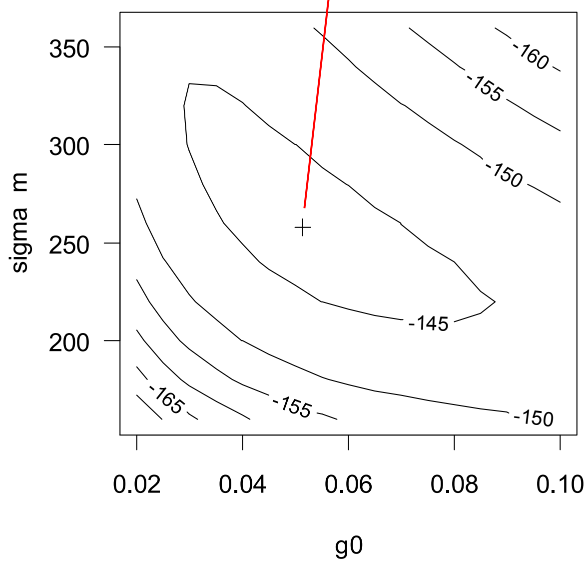
```
Fitted (real) parameters evaluated at base levels of covariates
       link    estimate  SE.estimate          lcl          ucl
D       log  0.02236315  0.006542529   0.01275268   0.03921609
g0    logit  0.05146938  0.021702334   0.02220028   0.11479676
sigma   log 257.90358775 44.727329279 184.04698278 361.39826673
```

model

```
Model             :  D~1 g0~1 sigma~1
Fixed (real)      :  none
Detection fn      :  halfnormal
Distribution      :  poisson
N parameters      :  3
Log likelihood    :  -144.0016
AIC               :  294.0033
AICc              :  295.5033
```

Poisson vs binomial *n*

maximum



LLsurface.secr (stoat.model.HN, c("g0", "sigma"),
xval = seq(0.02,0.10,0.005), yval = seq(160,360,20))

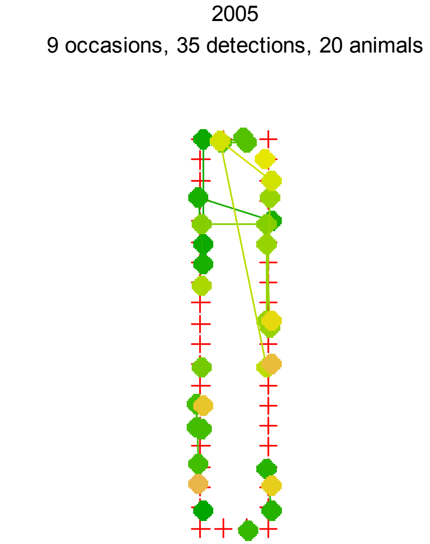
D held constant at ML estimate

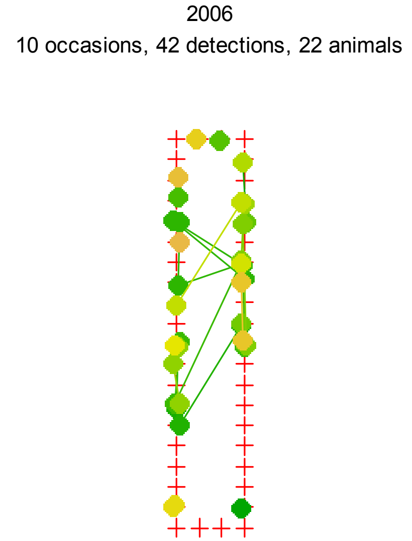
Ovenbirds at Patuxent Wildlife Refuge, MD

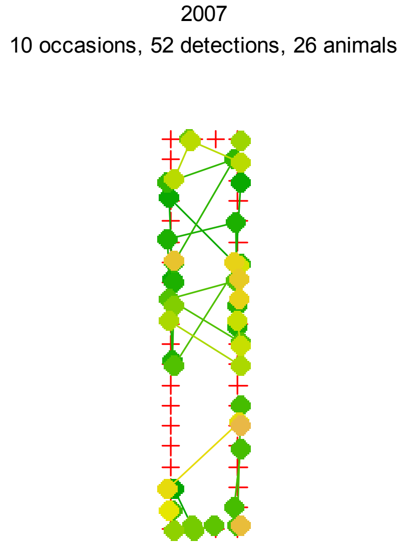
May/June
2005–2009

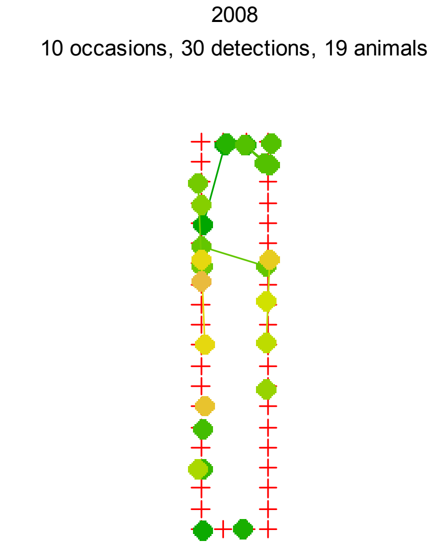
```
data(ovenCH)
counts(ovenCH)

$`M(t+1)`
      1  2  3  4  5  6  7  8  9 10 Total
2005  5 12 13 14 15 16 16 18 20 NA    20
2006  5  9 11 16 19 19 19 21 21 22    22
2007 12 15 16 18 20 20 22 23 25 26    26
2008  7  9 10 11 12 12 14 18 18 19    19
2009  4  7 11 13 13 14 14 15 16 16    16
```

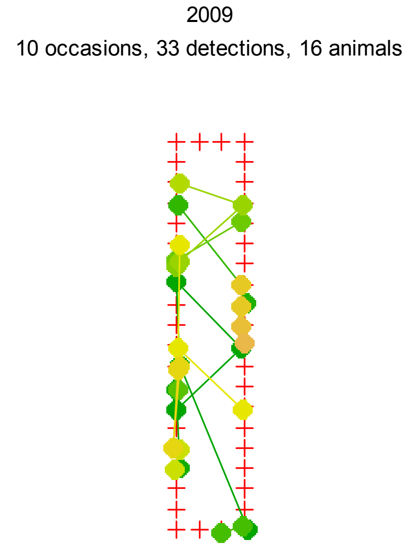
```
plot(ovenCH, gridlines = F, varycol = T, tracks = T)
```











```
secr.fit( capthist = ovenCH, model = list(g0 ~ 1), mask = ovenmask )
secr.fit( capthist = ovenCH, model = list(g0 ~ b), mask = ovenmask )
```

```
AIC(ovenbird.model.1, ovenbird.model.1b)

                             model   detectfn npar    logLik      AIC     AICc dAICc  AICwt
ovenbird.model.1b D~1 g0~b sigma~1 halfnormal    4 -927.5894 1863.179 1863.587 0.000 0.9219
ovenbird.model.1  D~1 g0~1 sigma~1 halfnormal    3 -931.1404 1868.281 1868.523 4.936 0.0781
```

```
collate(ovenbird.model.1, ovenbird.model.1b)[1,,,'D']

                   estimate SE.estimate       lcl      ucl
ovenbird.model.1  0.9105891   0.1256642 0.6956767 1.191893
ovenbird.model.1b 0.7070765   0.1067522 0.5268391 0.948975
```

A couple of SECR myths:

1. “SECR is for density D, CR is for population size N”

2. “SECR estimates are imprecise”

Population size in a defined area from SECR model

region.N {secr} R Documentation

Population Size

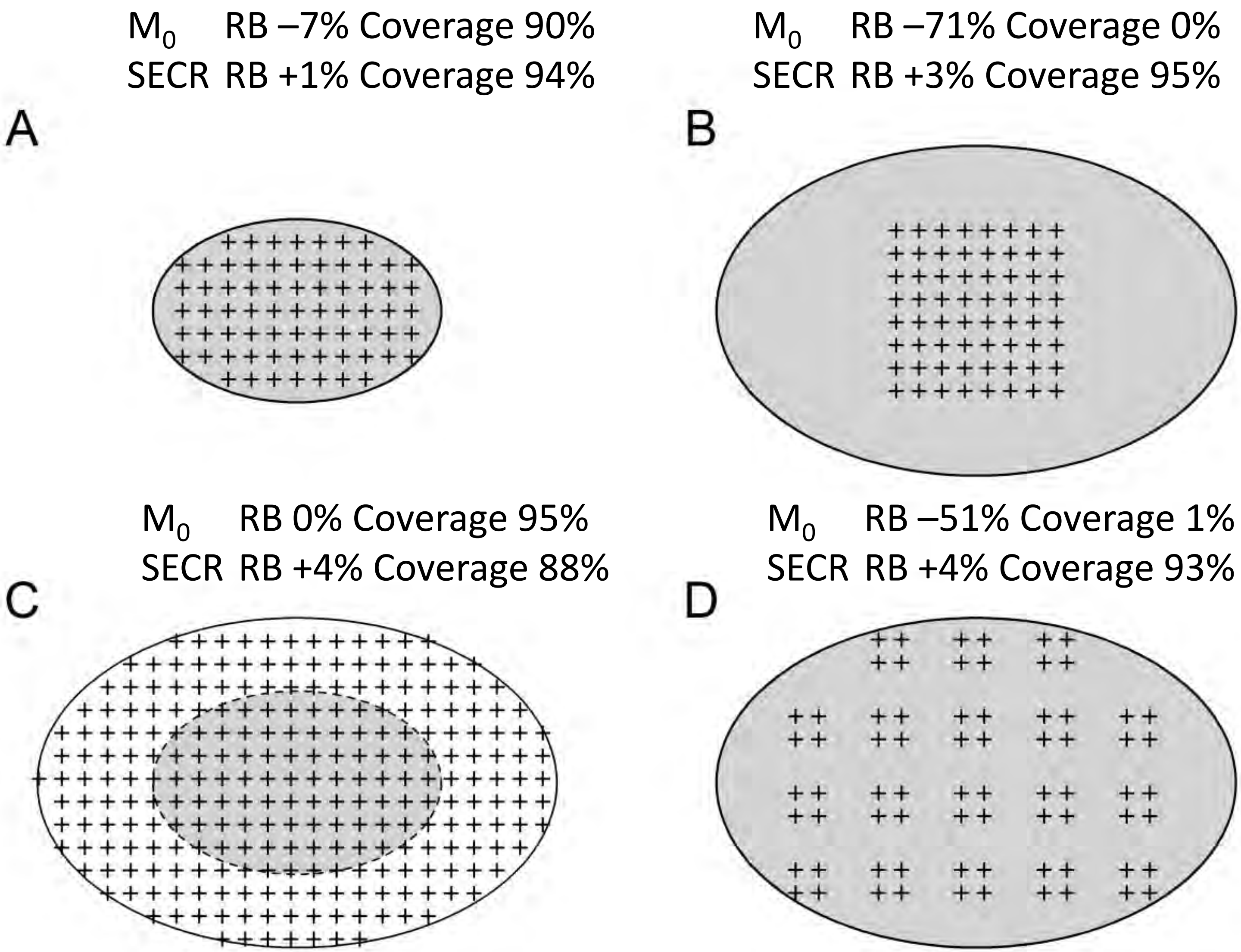
Description

Estimate the expected and realised populations in a region, using a fitted spatially explicit capture–recapture model. Density is assumed to follow an inhomogeneous Poisson process in two dimensions. Expected *N* is the volume under a fitted density surface; realised *N* is the number of individuals within the region for the current realisation of the process (cf Johnson et al. 2010; see Note).

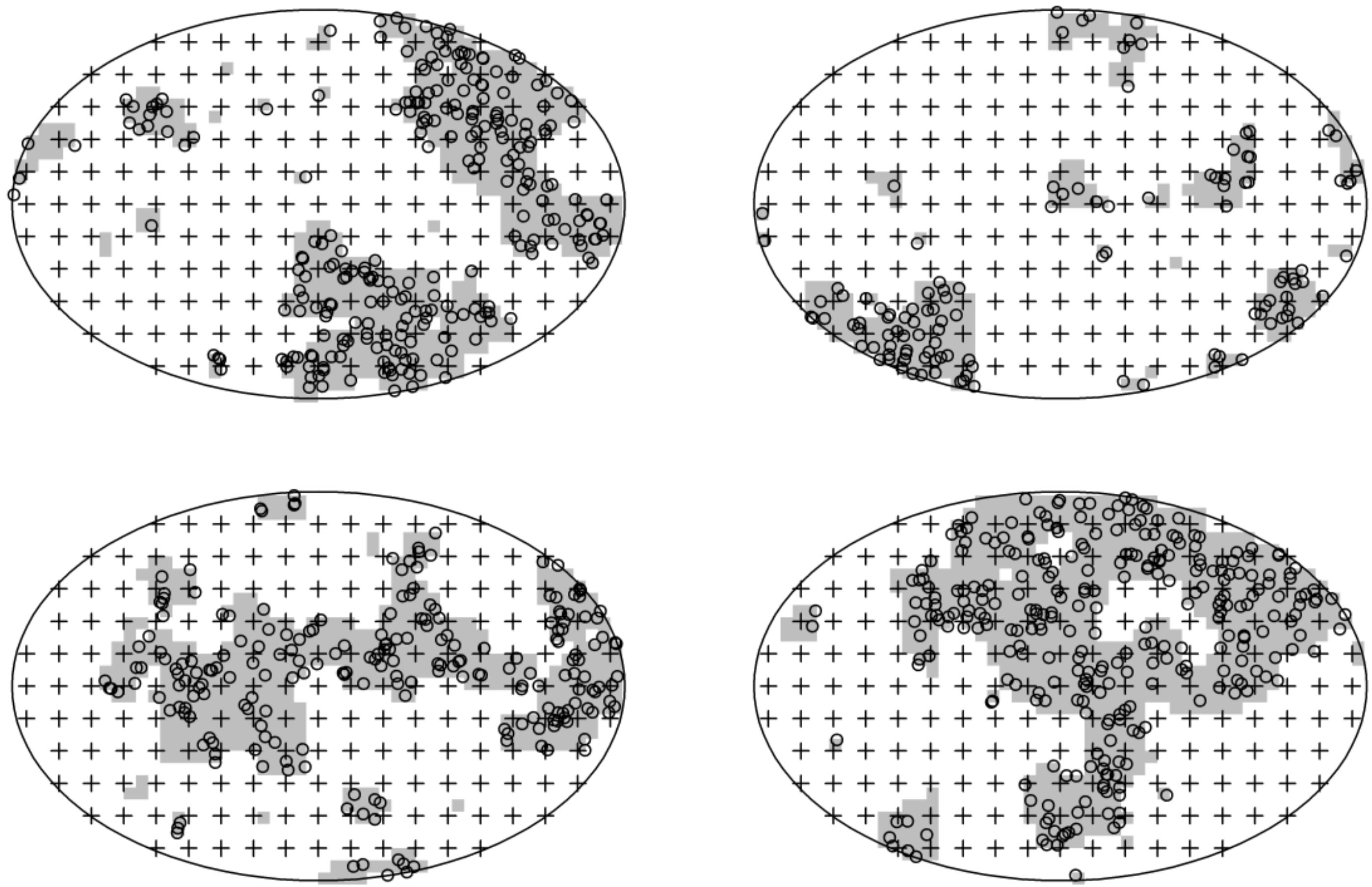
Usage

```
region.N (object, region = NULL, spacing = NULL, session = NULL,
    group = NULL, se.N = TRUE, alpha = 0.05, loginterval = TRUE,
    keep.region = FALSE, nlowerbound = TRUE, RN.method = 'poisson')
```

Nonspatial vs spatial estimates of population size – Efford & Fewster in review



Comparing non-spatial and spatial estimates of N for random landscapes

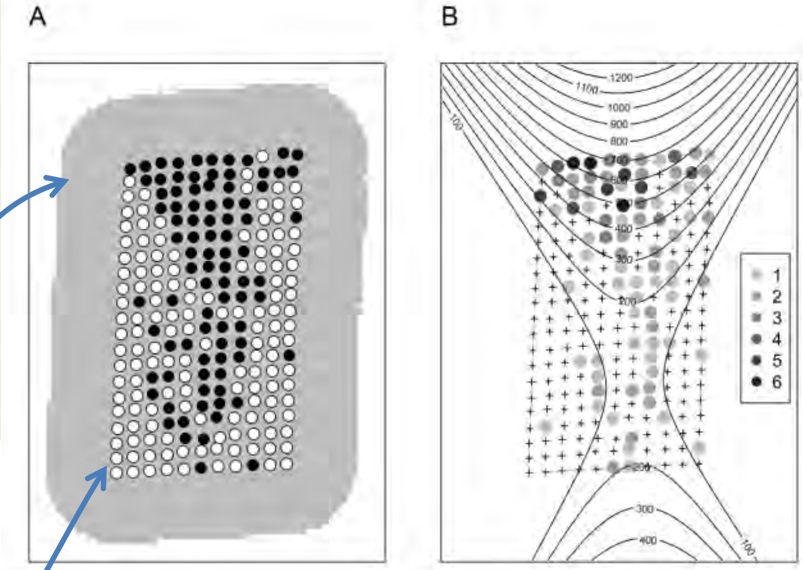


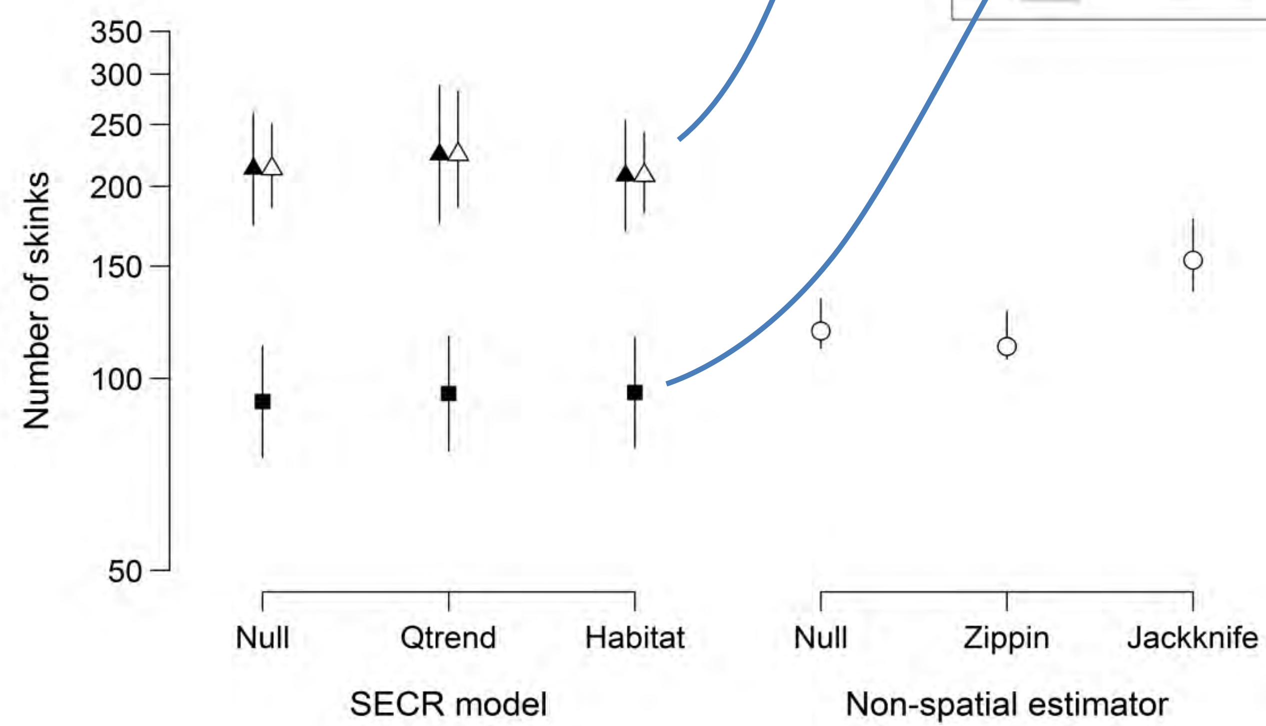
Summary of simulation results:

M_0 RB -3% Coverage 93%

SECR RB 0% Coverage 94%

Population size with modeled density – estimates relate to a known area





“SECR estimates of N less precise than conventional ones”

Scenario	Estimator	RB	RSE	Coverage
A	M_0 null	–0.068 (.003)	0.090 (.001)	0.903
	M_b Zippin	–0.036 (.009)	0.178 (.004)	0.906
	M_h jackknife	+0.095 (.005)	0.112 (.001)	0.779
	SECR $\hat{N}$	+0.010 (.003)	0.098 (.001)	0.940
	SECR $\hat{\mu}$	+0.004 (.005)	0.152 (.000)	0.945

Not much difference